Herding Concepts: The Contextual Validation of Social Agent Models[*]

David L. Sallach
Argonne National Laboratory
University of Chicago

0. ABSTRACT

If social models are to be used to inform policies, their validation is a vital topic. However, the validation techniques of the natural sciences are not sufficient to validate social models. The present paper considers the concept of contextual validation and how it can address these additional requirements. The discussion illustrates how contextual validation is applicable to both substantive and computational social science research, and how propensity theory can be used to achieve its benefits.

1. INTRODUCTION

The promise of social agent modeling, and computational social science (CSS), is widely recognized (Gilbert & Troitzsch 2005; Epstein 2006; Miller & Page 2007; North & Macal 2007; Lin & Carley 2010). The capability of representing social processes at multiple scales, at an arbitrary level of detail, is required, and enables the construction of rich, interwoven models with the potential to provide a more effective map of historical and policy-oriented dynamics.

However, the construction of complex models increases the requirement for rigorous validation. While intuition may provide a plausible means of assessment for simple and/or notional models, multiple interacting processes require a higher standard of validation.

In addition to complexity, *per se*, social models manifest a number of domain-specific validation challenges, including high-dimensionality, cultural indexicality, endogeneity, the lack of accessibility of actor motives and intent, diverse non-linear transitions and contextual effects (Sallach 2010). In the natural sciences, validation is primarily a focal process with model outputs being directly compared with relevant empirical data. While focal validation remains relevant in the social sciences (cf., Hartley & Starr 2010), these domain-specific challenges suggest that additional validation approaches may be required. One such approach, considered in the present discussion, is *contextual validation*. The focus of this approach is upon the scope, resilience and limits of the model, characteristics that help determine potential contribution of the modeling effort.

Succeeding sections of this paper explore contextual validation (section 2) including illustrations based on a sequence of theories of revolution scenarios (and the empirical exemplars from which they arise), demonstrate how propensity networks can serve as a foundation for contextual validation (section 3) and, briefly, conclude (section 4).

[*] This paper was written as part of the Model Evaluation, Selection and Application (MESA) project, funded by the Office of Naval Research (ONR), contract # DE-AC02-06CH11357. I am grateful for feedback received from colleagues at Lockheed-Martin, Lustick Consulting and Argonne National Laboratory.

2. CONTEXTUAL VALIDATION

It has often been observed that the maturity of natural science validation provides a valuable resource that can advance validation techniques within the social sciences (Hartley & Starr 2010; McNamara, *et al.,* 2010). While such transference between domains certainly can make a contribution, the special challenges of the social sciences require additional techniques. The one introduced in the present discussion, contextual validation, concerns the scope of applicability, permeability of boundaries and limits of boundaries that may (or may not) characterize the models in question. Successful resolution of such issues can significantly increase the clarity and confidence of computational models.

Challenges of Social Validation. Social phenomena significant enough to model occur within high-dimensional spaces that defy determinate and/or predictive modeling. It is possible that, for any specific problem type (war, insurgency, state collapse, etc.), all of the relevant events in recorded history may constitute only a small sliver of the possible events that could occur. From the perspective of probability theory, the available *sample* of events is not and cannot be representative of the *population* of possible events.

To the extent that this pattern turns out to be accurate, it presents a challenge to social modeling that cannot be addressed through validation alone. What a validation process can do is: alert modelers and stakeholders alike as to prospective risks associated with a model, and how those risks might be mitigated.

From a somewhat different perspective, consider several possible cases that involve model scope or boundaries:

- What would be the effect on an insurgency if a major epidemic broke out, including in the region of the insurgency?
- In the latter stage of a cultural transformation, a social revolution erupts with mutually perturbing consequences.
- What would be the effect of a mass campaign of civil disobedience, if the nation in which it was occurring began to engage in a major war?

These are only a few generic examples, but they illustrate how, even with a strong model of a primary focus (insurgency, cultural change, social movement) the expected results might be significantly altered. From a policy perspective, an important aspect of validation is to be able to specify circumstances under which the model can no longer provide credible results.[1] Model efficacy may also be undermined by smaller parallel developments, especially if and when they occur in combination. All such potential interactions are what necessitate the development and refinement of contextual validation.

To further illustrate the type of issues that contextual validation is designed to address, consider a sequence of conceptual innovations in the theory (and, thus, potential models) of social revolution.

---

[1] The emphasis of the present discussion is on the assessment of generic models. It is also possible to design empirical models, of course; however, such models will lack general utility and are likely to be ill-prepared for unanticipated developments. Thus, the focus of contextual validation addresses the robustness of models at varying degrees of generality.

Theories of Revolution. In the twentieth century, the theory of social revolution underwent significant development through the work, *inter alia*, of Edwards, Brinton, Skocpol, Sewell and Goldstone. Edwards (1927) and Brinton (1965) introduced comparative case studies as a means of developing a generalized stage theory of revolution, where the source cases were the English Civil War (1642), and the American (1776), French (1789) and Russian (1917) revolutions, with emphasis on the last two. Their major stages were instability and upheaval, the catalytic onset, the supersession of the moderates, the reign of terror/virtue, followed by retrenchment (e.g., Thermidor).

Such a script sequence, while sometimes suggestive, is not sufficiently generalized to be able to effectively provide the basis for a model of a single, open-ended revolutionary process. Not only is the cycle overly scripted, but it also ignores structural effects. An early correction was provided by Johnson (1966), who took an early step toward structuralism. More specifically, Johnson conceived effective social functioning as being based on ongoing equilibration processes. However, when societies are insufficiently responsive, a chronic disequilibrium develops that allows revolutionary leaders and movements to play a reequilibrationary role.

A deeper structural representation of the revolutionary process was provided by Skocpol (1979). Regarding the case studies upon which she bases her qualitative model, relative to Brinton, Skocpol drops the English and American revolutions, while adding the Chinese revolution (1949). More importantly, she replaces the stage theory with a structural theory of social revolution based on a conjunction of events that were, themselves, each independently caused. These events included state crises, popular uprisings and elite actions. As they unfolded, these events often were shaped by international forces that impinge on the states in question. In response to these events, the state and other elite actors found themselves constrained by the initial crisis (typically financial, often deepened by the exigencies of war) and therefore increasingly susceptible to the revolutionary challenges. The role for quasi-independent social actors, and the historically unique forms and sequencing of events made the theory more supple, and better able to represent diverse scenarios than prior theories.

These strengths notwithstanding, one of the important limitations of Skocpol's approach was the absence of cultural factors, as the Iranian revolution of 1979 dramatically illustrated. Skocpol noticed this omission herself (1982), but arguably in a narrow way, immediately identifying the structural locations of ideological groups, thereby integrating them into her conceptual framework. Sewell (1985) criticizes her cultural shortcoming and then, relying only on the history of the French revolution, demonstrates in some detail the richness of dynamic ideological influences. However, Skocpol (1985) critiques Sewell's treatment of ideology as being too synchronous and overly reified. The result is an edifying exchange, but a reasonable synthesis that incorporates ideology is left for later scholars.

Perhaps the fullest theoretical synthesis is provided by Goldstone (1991) who draws upon four early modern examples (Turkey 1629; China 1630; England 1642; France 1789) to broaden and deepen the theory of social revolutions. Specifically, he makes a clear distinction between the origin and the outcome of a social revolution. The former is governed by demographic and structural factors, while the latter is shaped by cultural and ideological influences that are, themselves, dynamically emergent.[2]

[2] Apter & Saich (1998) further developed the role of cultural factors by showing how social movements themselves give rise to a discourse community, the spread of which is a critical process within a revolutionary movement.

If it were assumed that models had been constructed that reflected each significant theoretical innovation in this history, the prospective contribution of contextual validation becomes clear. Specifically, we can concentrate on four conceptual transitions: 1) stage theory to social system equilibria, 2) equilibria to structural conjunction, and 3) narrow structures to demographics plus structure coupled with culture and ideological dynamics.

Each of these transitions can be viewed in terms of context validation, which can manifest in one of two ways. First, the theoretical history described in this section can be regarded as a research history (as it is), in which case contextual validation services can ensure that computational models of revolution reflect the most relevant theoretical precepts. Second, if we regard the same theoretical history as having given rise to correlative models, as it surely will in the future, contextual validation will then serve to assess the adequacy of existing theory, and explore how it might be usefully broadened, deepened or otherwise improved. That is, design-stage and contextual validation will contribute to, and potentially expedite, the type of theoretical progress summarized here.

While these theoretical transformations are by no means the only ones possible, they do illustrate model broadening and deepening benefits from contextual validation as a means of assessing the relative strengths assumed, and implicit weaknesses, in the modeling process. However, for the contributions of contextual modeling to be fully realized, it is essential to establish the foundations of its emerging framework. This topic is addressed in the next section.

3. PROPENSITY NETWORKS

To this point, the focus has been on broad models capable of providing a rich context for policy-oriented decisions, and how best to validate them. Contextual validation has been proffered as a technique that facilitates this process. However, for contextual validation to provide its full contribution, it requires a rigorous foundation. It is to this process that the discussion now turns.

Propensity theory was developed to provide an alternate foundation for probability theory. Because decision problems require a specific-case focus, a frequentist interpretation severs the links between probability and decision theories (Albert 2007). The other primary interpretation of probability theory, Bayesian subjectivism, makes the determination of probabilities appear arbitrary. As a result, propensity theory provides a uniquely coherent link between publically accessible single-case probabilities and the decision process.

Propensity is a theoretical concept (Bartelborth 2011:367-371) that serves as a generalized, indeterministic theory of causality (Albert 2007:588). The scientific definition and use of propensity fields is based on a conceptual innovation of Karl Popper, the renowned philosopher of science (1990; 1959). In contrast to frequency interpretations of statistics, using examples from physics and gambling, *inter alia,* Popper maintained that sequences of outcomes can be best understood as a product of their *generative conditions*. Loaded dice, for example, depart from a uniform distribution of outcomes (between 1 and 6 for each die) whether they are rolled once or 10,000 times. Hence, he argued, probability is not a property of the sequence of rolls, but of the generating conditions.

Propensity fields can be used as a primary means of expressing, integrating and assessing diverse substantive models to which the decision framework is applied. Propensity fields are thus capable of expressing probabilistic (as well as determinate) relationships across a wide range of disciplines (Sallach, *et al.,* 2008).

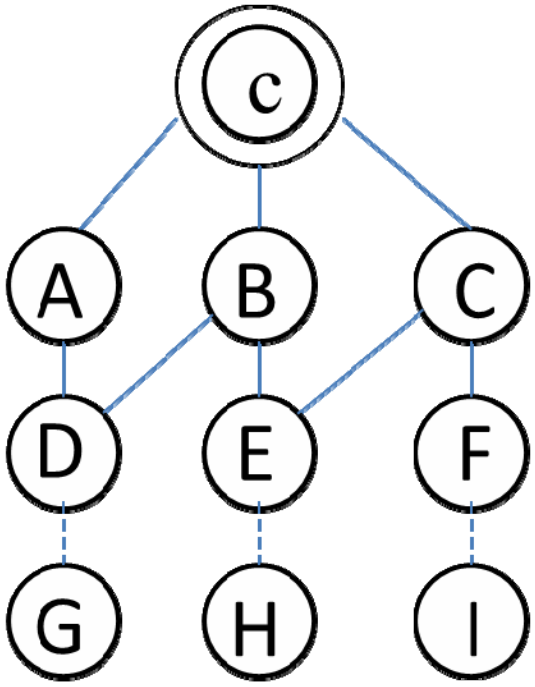


Figure 1. Stylized Propensity Network

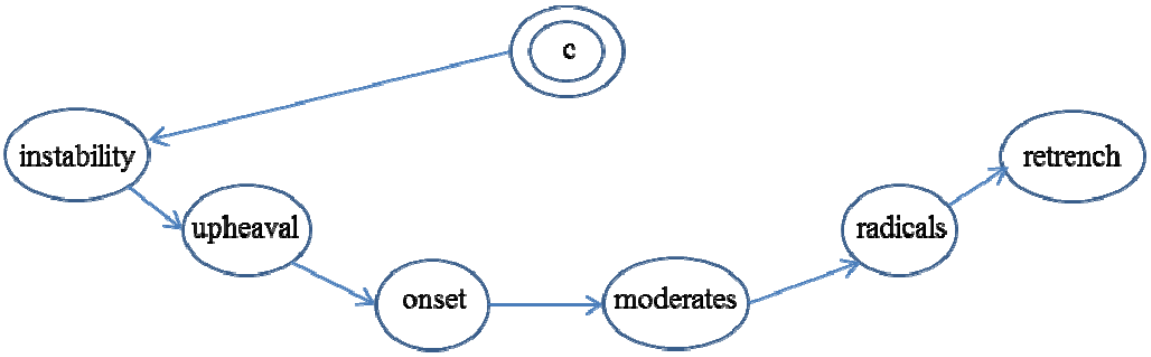
Propensity-theoretic patterns can be effectively represented in graph form, as a propensity tree or network (cf., Albert 2007). As an introduction to the conventions of this formalism, Figure 1 shows an illustrative propensity network. Any double-circle node represents a source of causation within the network. In this example, Conditions (c) represents all unspecified sources of causation.

A single-circle node represents a causal factor within the network. Predecessor nodes can exert an exclusive causal influence, as in C-F or a joint influence as in the blend of the A-D and B-D relations. Single-circle nodes may represent types or instances of values, depending on the level of specificity of the model in question.

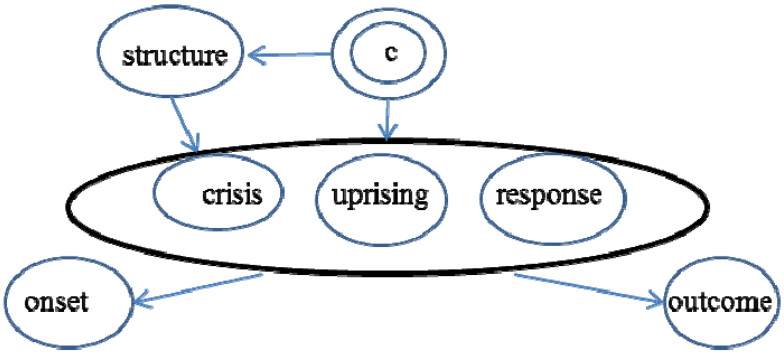
When used for purposes of contextual validation, dotted relations can be introduced. These links represent proxy factors. In this case, D represents a concept, while G represents available data that may serve as a proxy for that concept (and so too for factors H and I relative to E and F, respectively. These proxies may be imperfect indicators and, for validation purposes, it will be important that they be carefully assessed.

However, in this illustration, for purposes of contextual validation, the 'c' node plays a vital role. The adequacy of scope of a model may be determined by causes that have not been differentiated. In the context of the prior discussion of the evolution of theories of revolution, Figure 2 illustrates how illustrative theories (and, by extension, models based upon them) exhibit a higher level of context validity (and/or a more appropriate level for a particular type of problem) than earlier formulations provide.

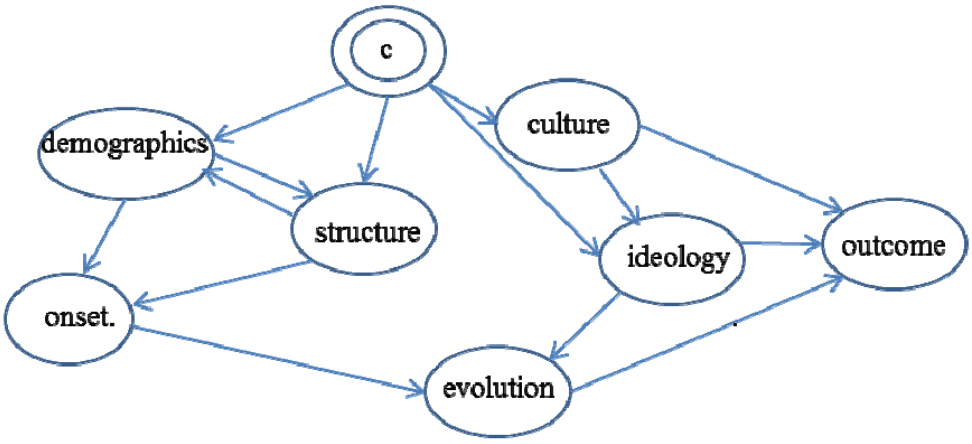
The first example in Figure 2a is a Brinton-type stage theory. The revolutionary onset is lost in undifferentiated circumstances, while the progression is in standard stages that need only be parameterized. The example in Figure 2b is a Skocpol-type structural theory in which social



(a) A Stage Model of Revolution



(b) A Structural Conjunction Model of Revolution



(c) A Synthetic Model of Revolution

Figure 2. Multiple Models of Revolution

structure and other generic conditions create the tripartite prerequisites for social revolution. However, origin and outcome are not differentiated, and cultural factors play no role. Figure 2c is a more sophisticated Goldstone-type theory in which origin and outcome have been distinguished and the causal factors that influence them have been differentiated as well.

In practice, each of the examples in Figure 2 would be developed in greater detail. The distinction between types and instances would be represented, the adequacy of data proxies would be assessed and quantitative propensity values would be specified. However, for the purposes of the present discussion, the comparative examples illustrate how contextual validation can be specified. The *ceteris paribus* state suggested by the source (c) node is also the source of undifferentiated but potentially significant factors or mechanisms.

The search for such mechanisms, which is abductive in nature, is neither universal nor monolithic. On the contrary, the contextual factors of interest are likely to be historically specific. However, their contingency does not diminish their relevance. Contextual events and process often operate to severely undermine the credibility of an otherwise useful model. It is for this reason that contextual validation must be refined as a working tool and utilized by modelers and stakeholders alike.

4. CONCLUSION

To the extent that social agent models are used for strategic or policy-oriented purposes, model validation is, and will remain, a vital issue. However, the mature validation techniques of the natural sciences are, by themselves, insufficient for the requirements of the social sciences. One important aspect of this challenge is the extent to which multiple simultaneous processes permeate each other and, thereby, interactively shape unique and unexpected outcomes.

The present discussion has suggested that the scope and limits of models are as important as the representation of their focal processes. Contextual validation has been proposed as a means of exploring and assessing model scope and limits in particular applications. The evolution of twentieth century theories of revolution has been used to illustrate how theoretical development has been driven by context shifts, and how propensity networks provide a method by which contextual validation can be implemented.

5. REFERENCES

Albert, Max. 2007. The propensity theory: A decision-theoretic restatement. *Synthese* 156:587-603.

Apter, David E. and Tony Saich. 1994. *Revolutionary Discourse in Mao's Republic*. Cambridge, MA: Harvard University Press.

Bartleborth, Thomas. 2011. Propensities and transcendental assumptions. *Erkenn* 74:363-381.

Brinton, Crane. 1965 [1938]. *The Anatomy of Revolution.* New York: Vintage.

Edwards, Lyford P. 1927. *The Natural History of Revolution.* Chicago: University of Chicago Press.

Epstein, Joshua M. 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling.* Princeton, NJ: Princeton University Press.

Gilbert, Nigel & Klaus G. Troitzsch. 2005. *Simulation for the Social Scientist.* New York: Open University Press.

Gintis, Herbert. 2009. *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences.* Princeton, NJ: Princeton University Press.

Goldstone, Jack A. 1991. *Revolution and Rebellion in the Early Modern World.* Berkeley, CA: University of California Press.

Hartley, Dean & Stuart Starr. 2010. Verification and validation. Pp. 311-336 in A. Kott & G. Citrenbaum, eds., *Estimating Impact: A Handbook of Computational Methods and Models for Anticipating Economic, Social, and Security Effects in International Interventions.* New York: Springer.

Johnson, Chalmers. 1966. *Revolutionary Change.* Standford, CA: Stanford University Press.

Lin, Zhiang & Kathleen Carley. 2010. *Designing Stress Resistant Organizations: Computational Theorizing and Crisis Applications.* New York: Springer.

McNamara, L.A., T.G. Trucano, C. Geisler & J.L. Perry, eds. 2010. *Challenges in Computational Social Modeling and Simulation for National Security Decision-Making.* Fort Belvoir, VA: Defense Threat Reduction Agency.

Miller, John H. & Scott E. Page. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life.* Princeton, NJ: Princeton University Press.

Moshiri, Farrokh. 1991. Iran: Islamic Revolution Against Westernization. Pp. 116-135 in J.A. Goldstone, T.R. Gurr & F. Moshiri, eds., *Revolutions of the Late Twentieth Century.* Boulder, CO: Westview Press.

North, Michael J. & Charles M. Macal. 2007. *Managing Business Complexity: Discovering Strategic Solutions with Agent-Based Modeling and Simulation.* New York: Oxford University Press.

Popper, Karl. 1990. *A World of Propensities*. Bristol, UK: Thoemmes.

Popper, Karl R. 1959. "The propensity interpretation of probability." *British Journal for the Philosophy of Science* 10:25-42.

Sallach, David L. 2010. Social science model validation: Domain-specific challenges. In L.A. McNamara, T.G. Trucano, C. Geisler & J.L. Perry, eds., *Challenges in Computational Social Modeling and Simulation for National Security Decision-Making.* Fort Belvoir, VA: Defense Threat Reduction Agency.

Sallach, David L., Charles M. Macal, Jonathan Ozik & Michael J. North. 2008. SADF/*ChoiceWright*: A Decision Framework/Tool for Multiscale Social Models. World Congress on Social Simulation, Fairfax, VA.

Sewell, William H. 1985. Ideologies and social revolutions: Reflections on the French case. Journal of Modern History 57 (March):57-85.

Skocpol, Theda. 1979. *States and Social Revolutions.* New York: Cambridge University Press.

_____. 1982. Rentier state and Shi'a Islam in the Iranian revolution. *Theory and Society* 11 (May):265-283.

_____. 1985. Cultural idioms and political ideologies in the revolutionary reconstruction of state power. A rejoinder to Sewell. *Journal of Modern History* 57 (March):86-96.