

---

# CHAPTER 18

## QUANTITATIVE MODELS FROM QUALITATIVE DATA: CASE STUDIES IN AGENT-BASED SOCIO-POLITICAL MODELING

---

**David S. Dixon**

*Many socio-economic policy, planning and assessment questions arise because not enough is known about their subjects. While inaccessibility and lack of hard data are the very challenges that may make a computer model invaluable, they are also reasons why many modeling and simulation applications are never undertaken. The author has found that qualitative agent-based models that are appropriately focused can prove surprisingly rich in quantitative data. Such models, accompanied by a thorough delineation of the applicable scope and context, have provided important insights into otherwise inscrutable systems. Building on early lessons learned in qualitative modeling (Dixon & Reynolds 2003), broader issues of qualitative modeling are explored. Case studies include historical research, negotiations, leadership succession, and coalition formation.*

**Dave Dixon** has been in software design and computer programming since 1975, with concentration on agent-based modeling since 1997. With the company he cofounded in 1998 he specialized in agent-based models of socio-political systems for non-governmental organizations, the military and the intelligence community. Mr. Dixon has an M.S. in Physics (thesis in astrophysics) and is presently a Ph.D. candidate in Natural Resource Economics at University of New Mexico.

Excerpted from *Complexity and Policy Analysis: Tools and Concepts for Designing Robust Policies in a Complex World*, ed. Linda Dennard, Kurt A. Richardson & Goktug Morcol, 323-336. Goodyear, AZ: ISCE Publishing, 2008. ISBN 978-0-9817032-2-0.

## Introduction

Policy-makers rely on analysts, and when a policy-maker refers to a *good analyst*, that typically means an analyst who is able to weave data into narrative without doing injustice to either. Computer models tend to produce output which taxes the storytelling skills of even the best analysts. Useful models of complex systems are complex systems themselves, and those who promote their use must first master the skill of turning their models into narrative and vice versa.

The interplay of data and narrative is especially problematic when modeling data-poor systems. This is the case in most socio-economic policy issues that are outside purely financial or demographic areas. It is certainly true of international arenas when considering closed governments, multi-national non-government organizations, or segments of populations or economies that are not well known even by their own governments. Often, notional models—on paper or in computers—may be the only way to assess these systems or plan for contingencies that involve them.

Historically, lack of data has been seen as a limit to computer modeling, and analysts have made careers out of carrying trusted mental models of these systems in their heads. These analysts have always been in shortage, are impossible to scale, don't yet hook directly to the internet, and eventually retire from the workforce. Computer models are not likely to replace those analysts, but they could serve to replicate their stored knowledge and insights, alleviating the shortage, scalability and connectivity problems.

In order to address both the input problem—turning narrative-based systems into quantitative models—as well as the output problem—communicating quantitative models to a narrative-based world—computer modeling and simulation must be embedded within a narrative context. This paper reports on four case studies for which qualitative computer models were developed. In three cases the models produced quantitative results that were then interpreted in narrative form. In the other case the goal was achieved through the interchange between narrative and model design before actually implementing the computer model.

## The Goal of Qualitative Modeling

Gell-Mann defines the complexity of a system as the “length of the shortest message that will describe a system ... employing language, knowledge and understanding that both parties share” (Gell-Mann, 1994). It follows, therefore, that a problem can be made less complex by finding—or synthesizing—the right context. It could be said that most scientific endeavors consist of the discovery or synthesis of descriptive contexts that reduce the complexity of a topic by reducing the length of its description.

One hurdle is that real-world complex systems are typically multidisciplinary. Thus, Gell-Mann's “language, knowledge and understanding that both parties share” may be difficult to attain in practice. For each specialized field of study—physics, medicine, political science, and so on—there is a specialized lexicon. From each lexicon has arisen a body of lore: narratives that provide con-

cise and illuminating descriptions of the problems of interest in the field. The challenge for modelers is to find a narrative framework that borrows from the constituent fields without being exclusive to any one of them.

Often, graphical representation becomes the focus that unifies both the problem-statement narrative and the solution narrative in a multi-disciplinary way. It is no coincidence that highly effective graphics are typically multidisciplinary (Tufte, 1983). Tables and diagrams were the principle data-gathering tools in capturing the problem-statement narratives in all four case studies presented here. In three of the four cases, the same tables and diagrams proved ideal media for turning results into narrative (modeling results in one case, simulation results in the other two). In the fourth case, the unforeseen outcome was that the underlying driver was dynamical, so that a time-series representation became the basis for the results narrative.

The appropriate modeling paradigm should follow naturally from the choice of narrative framework and graphical medium. A number of modeling approaches have met with success in social science simulation (Gilbert, 1999), one of which is agent-based modeling (ABM). The case studies in this paper all employed ABM, which is an especially natural representation of the causal links between actors (individuals, groups, nations, *etc*) and their actions. An effective ABM can help to tell the story behind quantitative data in some cases, and highlight where important data are missing in others.

## Lessons Learned in Qualitative Modeling

The four case studies will be referred to in subsequent paragraphs by these abbreviations:

- SMALL—a small-group decision-making model (historical model);
- AMNESTY—a negotiation model;
- SUCCESSION—a leadership-succession model (voting model);
- COALITION—a coalition-formation model.

Each model presented in this paper began with extensive interviews with one or more domain experts. Flexible ABM software made it easy to refine the models in subsequent meetings with the domain experts. Many of those refinements came from merely reviewing what was modeled from previous interviews. Often, when formalized in a model, conventional wisdom turned out to be at odds with direct observation.

Occasionally the modeling process alone led to useful insights and simulation became unnecessary (SMALL). In some cases, however, the phenomena of interest were, by their nature, dynamical (AMNESTY). In these cases, domain experts refined their models based on both review of the model and review of simulation outcomes. In the third and fourth cases (SUCCESSION and COALITION) the models themselves were static while the overall problem space was simulated.

The first step in these case studies was to identify appropriate subject matter experts (SMEs). The SMEs in these case studies were recognized authori-

ties in their areas. In the first step, the SMEs identified key players and high-priority topics. In two of the case studies (AMNESTY and COALITION) multiple SMEs were involved. For AMNESTY, the SMEs—who came from very different organizations—worked as a team. For these SMEs, who typically worked in isolation, the opportunity to work with others provided a tremendous creative boost. For COALITION, three SMEs were interviewed individually and three models were constructed. The three SMEs identified essentially the same high-priority topics, but very different key players. This is explored further in the detailed discussion of the model.

In the next step, the SMEs estimated the rank each key player would give each topic. In one case (SMALL) the key insight occurred before this step was reached. In another case (SUCCESSION), three of the players (voters) were also the candidates for succession. In this case, the SMEs estimated the voters' ranks for each candidate in addition to their rankings for the topics. In a third case (COALITION) the rankings were accompanied by estimates of each player's *flexibility*—how rigidly they were thought to hold their positions.

When multiple SMEs are involved, it is typical for them to have differing perspectives on how to define groups as key players, but there should be a consistent way to reconcile those differences (e.g., one definition is at a higher level of detail than the other.) It is common—and appropriate—to find considerable disagreement among SMEs on how each key player may rank each issue. These differences are typically attributed to the SMEs' background information (*tacit knowledge*) or to *bias*.

## Case Studies

The four case studies identified in this paper are illustrated here because each presented a specific kind of insight. Each case will be presented with some background information, some comments on the modeling process, and discussion of the class of insight presented.

### AMNESTY (2003)

Some problems come to analysis with volumes of data and no clue how they interrelate. Epidemiological studies early in an outbreak are often of this class, with vast minutiae regarding the infected individuals, their situations, behaviors, communities and so on. The epidemiologist then searches for a narrative that tells how an infectious agent connects a house painter in Cincinnati to a lawyer in Edinburgh. Most narratives come together as soon as the right data have been correlated, but in some cases the narrative is the product of an inspired guess, which then leads to the collection of crucial data.

A variation on this problem—where the unexplained data were model-generated—arose in this case study. It is a model of a government contending with multiple rebel groups. The SME team included government and military analysts as well as a specialist from a non-government organization<sup>1</sup>. The initial model was intended to explore what factors would promote negotiations be-

1. Adam Isaacson, Center for International Policy.

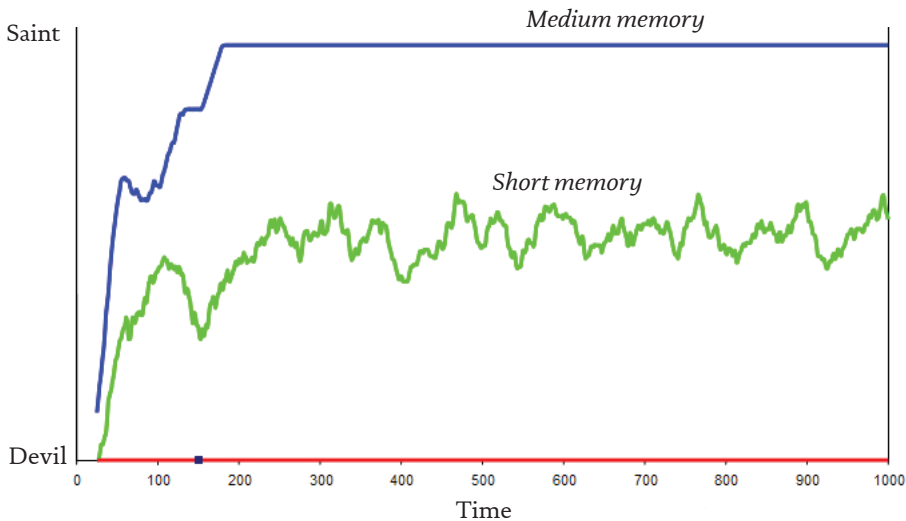


Figure 1 *Amnesty Model*

tween the rebels and the government and what factors might prevent them.

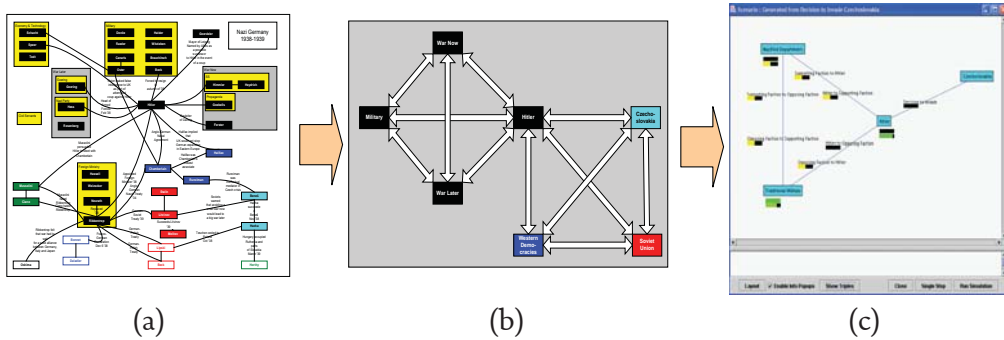
After initial specification, construction of the model, and multiple simulations and refinements, no values for the input parameters led to negotiations. Exploration of parameter space led to the realization that there was a hidden variable. The government in this model has long held, as a condition of negotiations, that a rebel group must refrain from *bad behavior* for a certain period of time. This had been taken as a given by the SMEs and modelers alike. Once this period of time was included as a variable, in the form of the government's *memory*, an exploration revealed that this one factor was the key driver to bringing the rebels into negotiations. Some discussion on this lead to the following narrative (see Figure 1):

*For infinite memory (nothing is ever forgiven) the rebels had no incentive to change their behavior (they were devils). For no memory (all is forgiven) the rebels had no incentive to change today, for there is always tomorrow. For some intermediate values of memory, the rebels did, eventually, clean up their actions (became saints) making negotiations possible. Finally, for some values between zero and intermediate, there were long periods of volatility, with many false starts but little or no progress toward negotiations. These reflected episodes of costly social turmoil, possibly worse than no negotiations at all.*

This study belongs to a class of problems for which there is a great deal of data, yet conventional wisdom and the experts miss a seemingly minor yet crucial variable. In this case, that variable was time.

### SMALL (2003)

Sometimes a modeling and simulation question leads to the right answer, other times to the right question. In these latter cases, as the narrative comes together, there are conspicuous gaps, resulting in a *collection requirement*—a place to look



**Figure 2** (a) Detailed model included every significant player in 1938. (b) First abstraction—Hitler, War Now, War Later, Military, and the external parties; Western Democracies, Soviet Union and Czechoslovakia. (c) Final model—Hitler, War Now, War Later, and Czechoslovakia.

for missing data. This case study, with Klaus Fischer<sup>2</sup> as SME, is an historical recreation model and was intended as a proof of concept. The model explored the decision by Germany to invade Czechoslovakia in 1938 (See Figure 2). A notional model was outlined as the SME recounted the events leading up to the invasion. In this case, that outline became the descriptive framework for this problem, because it provided the following narrative:

*In the years leading to the invasion, Hitler was a consensus builder, accomplishing this by carrot or stick, as appropriate. The evidence trail was there for every event except one: the military, which had opposed Hitler on the timing of the invasion, suddenly capitulated in August 1938, following the resignation of Army Chief of Staff General Ludwig Beck. What stick did Hitler use?*

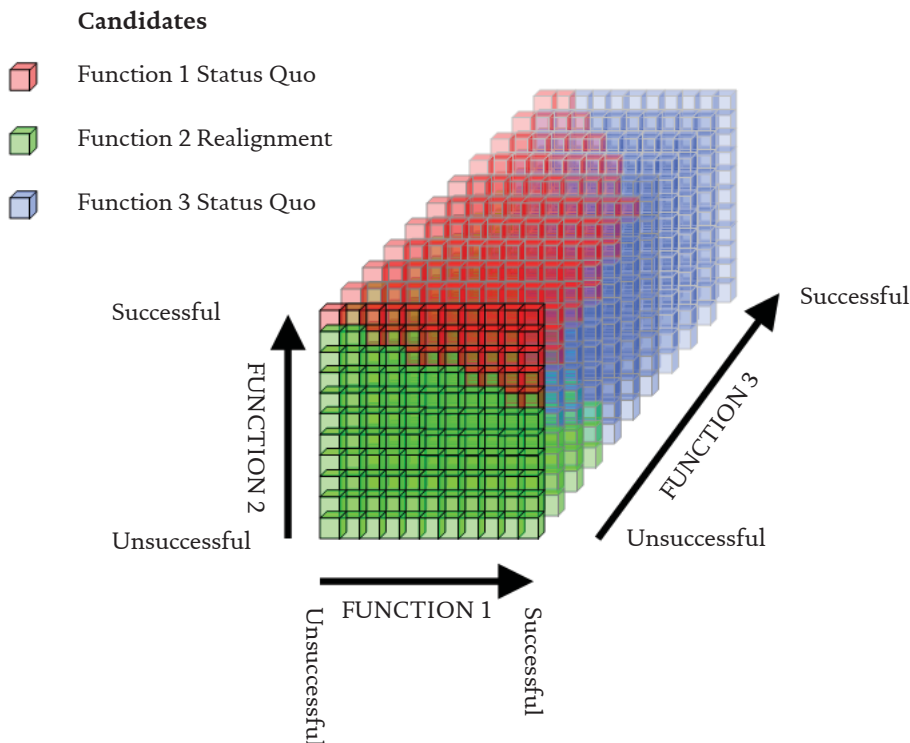
Further research by the SME uncovered that a secret proclamation in August 1938 diverted most of the military budget to the Nazi paramilitary wing, the *Waffen SS*, undermining the Army and General Beck.

This study belongs to a class of problems for which the systemization of data itself leads to the key insight. In this case, simulation became moot, but in others of this class, the insights gained in modeling provide completely new areas for simulation.

## SUCCESSION (2004)

This is another small group decision-making problem, but in this case the powerful central leader has left and the remaining members are deciding on a successor. The SME in this case is a government expert. Very little is known about members of the group aside from the three considered most likely successors, each of whom is associated with one of the three major functions of the organization. In a session with the SME, circles and arrows were drawn and tables were

2. Author of, among others, *Nazi Germany: A New History*, New York, Continuum, 1996. ISBN 0826409067.



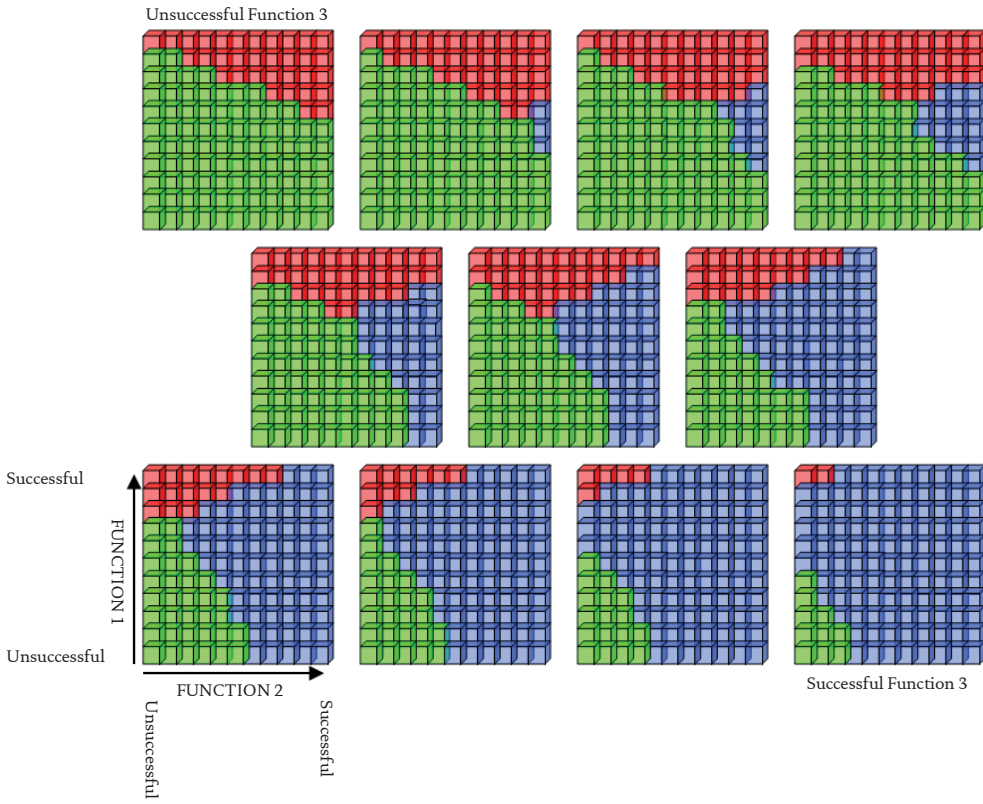
**Figure 3** *Three-dimensional Outcome Space for the Succession Model*

constructed to describe, numerically, the interactions between decision makers and the three candidates. These data became an initial, linear, static ABM from which the space of all possible outcomes was simulated. In fact, when there was a suggestion in the press that the group did choose a successor, a visit back to the model confirmed the choice.

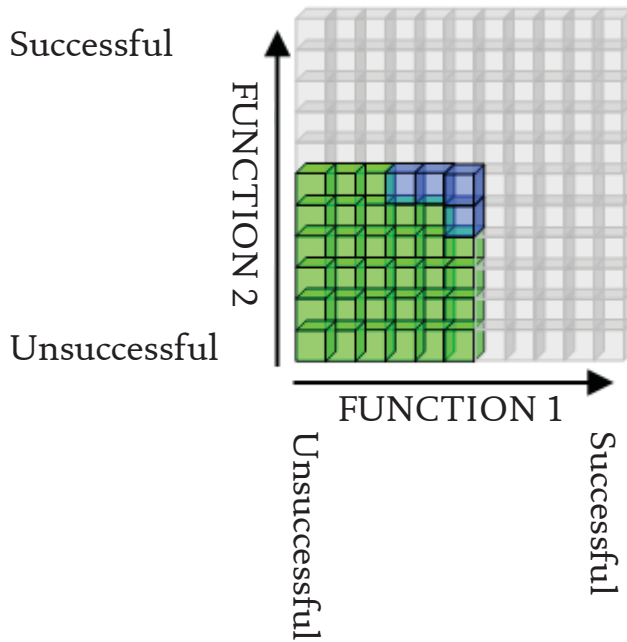
Figure 3 shows the three-dimensional outcome space. The three dimensions correspond to the three major functions mentioned above or, rather, to how well those functions are being performed. Each dimension is divided into eleven categories (ranks from zero to ten), resulting in the cube of cells shown in the figure. The color of a cell indicates the most likely successor under those circumstances.

The model was constructed many months before the succession appeared in the press, and at that time, the state of the three organizational functions was left unspecified. The model could have been used at any time simply by providing the best estimates of the status at that time, and yet the actual succession came as a surprise. When the news article appeared, and before referring back to the model, the modelers each guessed the outcome and all guesses were wrong. Yet a return to the model showed that the correct outcome was there all along. Was the model wrong?

Figure 4 shows the outcome space sliced along the Function 3 dimension. Part of the problem in interpretation came from *conventional wisdom* bias, and part with failing to believe the model. Conventional wisdom was that the strongman candidate (red) was the frontrunner. The modelers guessed the fi-



**Figure 4** Slices of Outcome Space for the Succession Model



**Figure 5** Mean over Function 3



nancier (blue). This perception came from the fact that Function 3 (finances) was known to be doing well, and the visual perception was that blue dominated at the high end of Function 3 (see the bottom right grid in Figure 4). Just a little further examination would have pointed to a very different conclusion.

The reality was that Function 1 and Function 2 had not been going well, and this information *was* known from the very beginning of the project. That puts the result in the lower-left quadrant of each of the slices in Figure 4. If we assume no knowledge whatsoever about Function 3 and take the mean, the result is shown in Figure 5.

That is, for this subset of outcome space, there is an eight out of nine, or 0.89, probability that green will be the successor. Green was, in fact, the successor identified in press reports.

This case belongs to a class in which insights are gained into a question not originally asked. Initially, this model was intended to be descriptive only: its predictive potential was realized only after the fact. The similarity to a *phase diagram* of the three dimensional outcome representation in Figure 3 introduced an entirely new descriptive framework and new ways to analyze the model, such as focusing on subsections of *phase space*, as in Figure 5.

### COALITION (2004)

In the year leading up to the January 2005 elections for the Transitional National Assembly of Iraq, there was wide divergence of opinion on who would be the key players. It was well known that there was a numerical majority of Shi'a, that Ayatollah Ali al Sistani was universally revered among them, and that a large number of their urban poor were also followers of Muqtada al Sadr. Unknown was how many Shi'a considered themselves religious versus secular or Iran-leaning versus nationalist, and what affect tribal affiliations (which often cross ethnic and religious divisions) might have. Similar unknowns applied to the large Sunni minority—the number of religious versus secular, the extent of continued Baath party allegiance, and the role of tribal affiliation. The large Kurdish minority, on the other hand, was well understood, but their desire for autonomy would make them unpredictable players in coalition forming. If the Shi'a were not able to establish a clear majority block—which was not foregone given the secular divisions—then the new government could be effective only with effective coalitions.

The goal of this case study was to explore the possibilities for coalitions within the Iraqi Transitional National Assembly (Dixon & Reynolds 2005). Three SMEs were interviewed. Each identified the high-priority topics as 1) the degree of governmental centralization, 2) the extent to which Islamic law (Sharia) would influence the legal system, 3) whether the Assembly would be clerical (like Iran) or secular, 4) the ongoing role of the U.S. in Iraq, and 5) the distribution of wealth (oil fields) and power (standing armies).

The first SME, Juan R. Cole<sup>3</sup>, divided the players into seven coarse groups (four Shi'a, two Sunni, and one Kurdish). Predictably, there was no common

3. Department of History, University of Michigan and author of *Informed Comment* (<http://www.juancole.com>).

ground and coalitions were found to be unlikely. The second SME, Amatzia Baram<sup>4</sup>, made similar groupings, but included miniscule fringe groups that turned out to straddle the larger divisions. These players had the potential of acting as catalysts in forming coalitions.

The third SME, Jeffrey White<sup>5</sup>, identified 14 players, dividing each Arab population (Shi'i and Sunni) into multiple religious groups plus secular and tribal groups. The two major Kurdish parties were identified individually, although they were estimated to hold nearly identical positions on all topics.

The potential for coalition formation is represented graphically in Figure 6 (Cole model), Figure 7 (Baram model), and Figure 8 (White model). The numbers in the cells represent relative distances between the groups. A light color represents similar positions (potential for coalition) while dark colors represent dissimilar positions (coalition unlikely). The diagonal is blackened, and the triangle below the diagonal shows the distances based on the SMEs rankings, while the upper triangle shows those distances diminished by the degree to which the parties are flexible.

Graphics like this one were produced for each of the first four topics (central government, Sharia, clerical rule, the U.S. role). The 14-player model also includes the distribution of wealth and power topic. The graphics shown here are a composite of the topics—the Euclidian distance in the 4-dimensional (or 5-dimensional) topic space.

Referring particularly to Figure 8, the checkerboard in the upper left represents the similar but not identical positions held by the Kurds and many Sunni. (Historical animosity between the groups is not portrayed.) The large white area in the center illustrates the great deal of accord among most of the Shi'i religious groups. It is interesting to note that the right-hand three columns, representing tribal, secular, and mainstream religious Shi'i groups, exhibit no strong differences with any group.

This is the narrative associated with Figure 6: The differences between the key players in Iraq are great and coalition government is extremely unlikely.

This is the narrative associated with Figure 7: Small groups may serve as catalysts in the formation of broader coalitions, similar to the student uprisings during the Iranian revolution.

This is the narrative associated with Figure 8: The religious Shi'a will form a stable block, the extent of which will depend on the inclusion of mainstream Shi'a, including al Sistani, who could bring secular and tribal factors into the coalition under the right circumstances. A coalition among Kurds and Sunni is possible, based on positions on topics, which could counter the numerical advantages of the Shi'a.

Note that these models were formed under the assumptions that the January 2005 elections would see participation from all communities. This did not turn out to be the case when the Sunni mostly boycotted the election, a situ-

---

4. Department of Middle Eastern History, University of Haifa, and, at the time of the study, Senior Fellow at the United States Institute of Peace.

5. Washington Institute for Near East Policy and former analyst with the Defense Intelligence Agency.

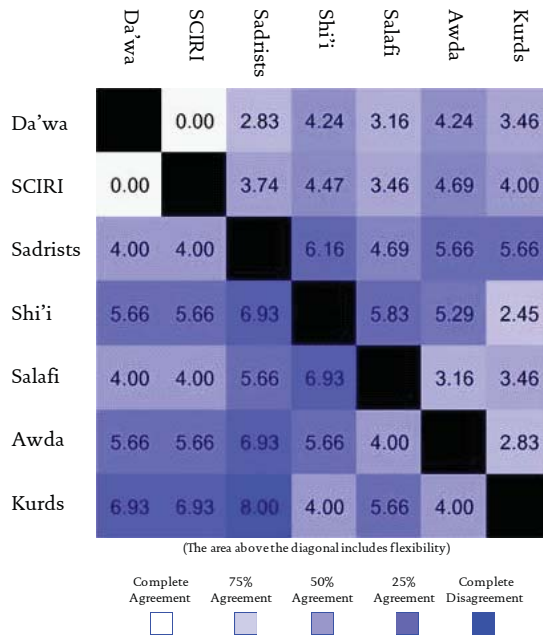


Figure 6 Cole (7-player) Iraq Model

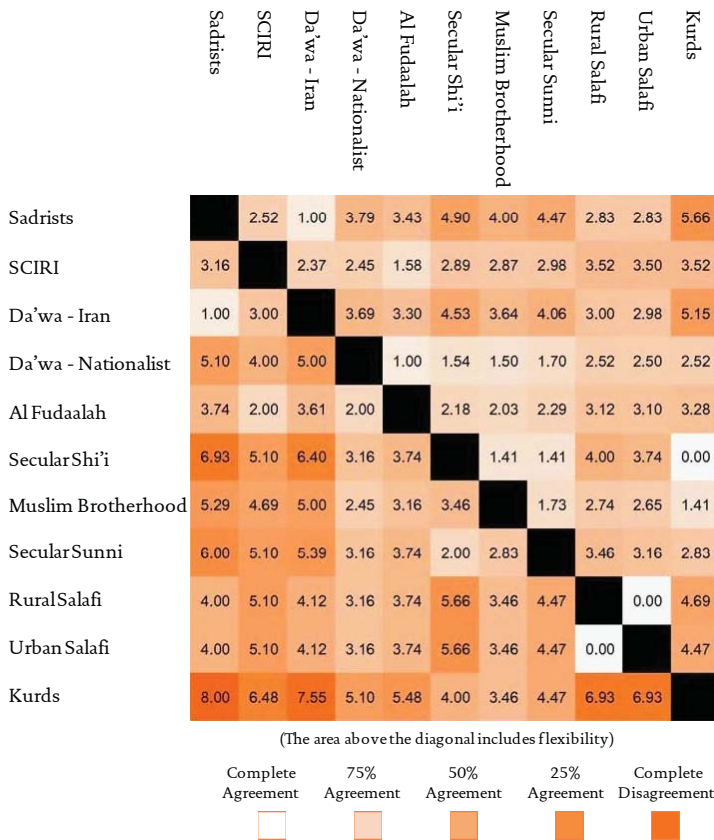


Figure 7 Baram (11-player) Iraq Model

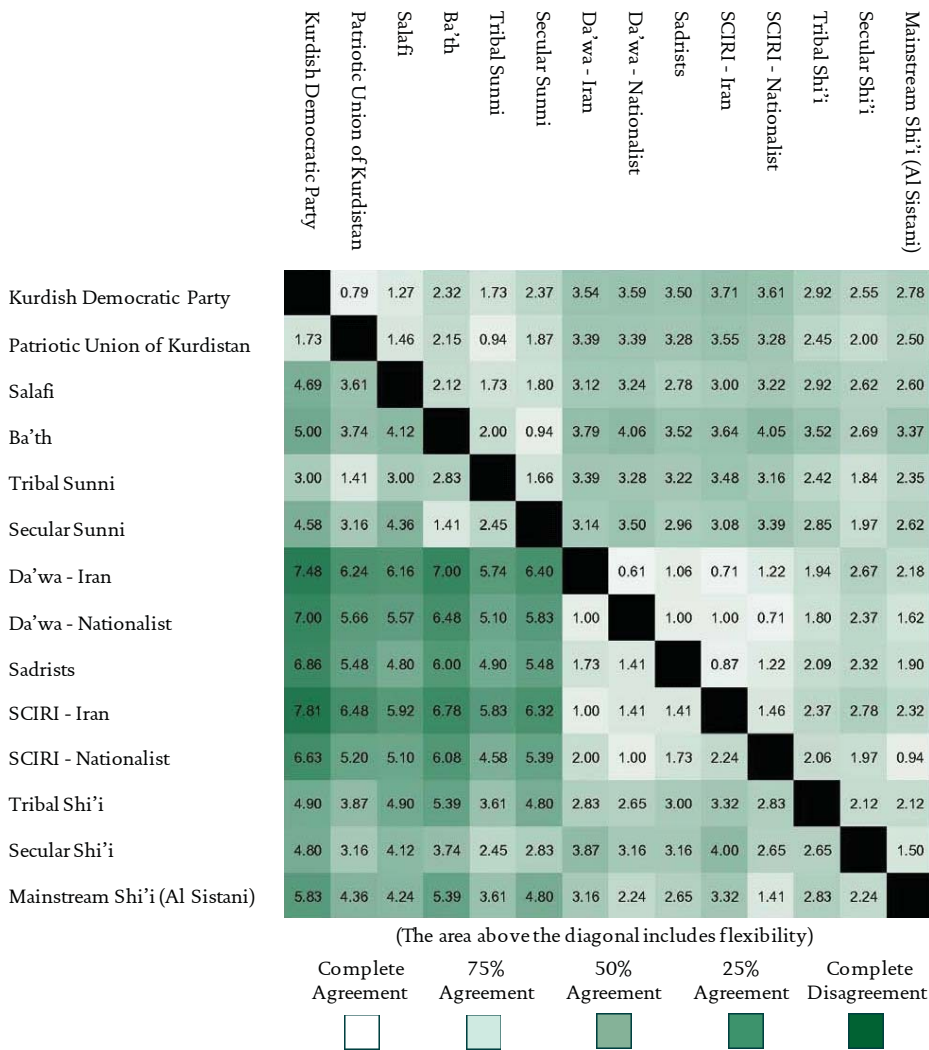


Figure 8 White (14-player) Iraq model

ation foreseen in a 2004 voting model developed by Loring White and Jeffrey White<sup>6</sup>.

In the January 2005 election, a broad religious Shi'a coalition, the United Iraqi Alliance, known by some Iraqis as "al-Sistani's list" (Wikipedia-UIA) took slightly more than 48% of the vote, the secular Shi'a party Iraqi List took nearly 14%, with the Kurdish coalition taking almost 26%, leaving slightly more than 12% of the vote to the other 108 parties (Wikipedia-Results).

This case belongs to the class of problems from which there is a great deal to learn about modelers and the modeling process, independent of the success of the model. The intelligence community, in particular, continues to seek insights into the roles of tacit knowledge and bias (Scholtz, *et al.*, 2005).

6. Personal Communication.

## Conclusion

Qualitative modeling and simulation have an important role in the analysis of complex policy issues. The success of modelers will continue to rest on their ability to embed computer models within narrative frameworks that both promote the formulation of effective models and the communication of results. The studies presented here are examples of four classes of insights that come from modeling complex social systems:

1. Despite large amounts of data, many problems may be driven by factors not captured by those data (stealth variables). Timing is often one such variable;
2. Much can be learned from the mere act of formally structuring existing knowledge;
3. A well constructed model may contain insights into (as yet) unforeseen issues. Hint: never throw away a model;
4. Some models may tell you more about the modelers and modeling process than the issue at hand. This is not a bad thing.

## Acknowledgements

This material is based upon work funded in whole or in part by the U.S. Government and any opinions, findings, conclusions, or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the U.S. Government. The author gratefully acknowledges the support of the Disruptive Technologies Office (DTO) for the case studies in this paper. The author is grateful, also, to the Department of Defense, which also supported two of the case studies. The case studies, and much of this paper, were completed while the author was affiliated with Least Squares Software, Inc. (LSS). The author acknowledges the significant contributions of William N. Reynolds of LSS, who collaborated on the case studies. Also, thanks to reviewers Cristina Pongiglione, David C. Earnest, and Cletus Moobela for their helpful suggestions. Finally, many thanks to Kurt Richardson for his tremendous efforts in organizing the conference and to Linda Dennard for hosting it.

## References

- Dixon, D. and Reynolds, W. (2003). "The BASP agent-based modeling framework: Applications, scenarios and lessons learned," *Proceedings of the 36<sup>th</sup> Hawaii International Conference on System Sciences*, Big Island, Hawai'i. 6-9 January, ISBN 0769518745.
- Dixon, D. and Reynolds, W. (2005). "Visualizing the political landscape," *International Conference on Intelligence Analysis*, McLean, Virginia, 2-6 May, <http://www.least-squares.com/papers/icia2005.pdf>.
- Gell-Mann, M. (1994). *The Quark and the Jaguar: Adventures in the Simple and Complex*, ISBN 0716725819.
- Gilbert, N. and Troitzsch, K. (1999). *Simulation for the Social Scientist*, ISBN 0335197442
- Sholtz, J., Crosby, M. and Ward, P. (2005). "Integrating humans with intelligent technologies: Merging theories of collaborative intelligence and expert cognition," *Pro-*

ceedings of the 38<sup>th</sup> Hawaii International Conference on System Sciences, Big Island, Hawaii. 3-6 January, ISBN 0769522688.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*, ISBN 0961392142.

The whole book, but the Introduction and p. 177 in particular.

Wikipedi-UIA, [http://en.wikipedia.org/wiki/United\\_Iraqi\\_Alliance](http://en.wikipedia.org/wiki/United_Iraqi_Alliance) .

Wikipedia-Results, [http://en.wikipedia.org/wiki/Iraqi\\_legislative\\_election\\_results%2C\\_2005](http://en.wikipedia.org/wiki/Iraqi_legislative_election_results%2C_2005) .