

AGENT-BASED MODELING IN INTELLIGENCE ANALYSIS

by

Aaron B. Frank
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Computational Social Science

Committee:

_____ Chair of Committee

_____ Director of Graduate Studies

_____ Director, Krasnow Institute for
Advanced Study

Date: _____ Fall Semester 2012
George Mason University
Fairfax, VA

Agent-Based Modeling in Intelligence Analysis

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at George Mason University

By

Aaron Benjamin Frank
Master of Arts
Georgetown University, 1999
Bachelor of Arts
Boston University, 1997

Director: Robert Axtell, Department Chair
Computational Social Science

Fall Semester 2012
George Mason University
Fairfax, VA

Copyright © 2012, Aaron B. Frank
All Rights Reserved

Dedication

For Desmond Saunders-Newton, an exceptional mentor and friend whose influence is felt every day.

Acknowledgments

I am deeply indebted to many people who offered their assistance, advice, experience and encouragement over many years. I would like to thank my chair, Robert Axtell, and my committee, Steven Bankes, Matthew Hendrey, Timothy Hoyt and Hilton Root for their patience in enduring and guiding me through multiple changes in scope, structure and topics. Likewise, James Olds and the Krasnow Institute for Advanced Study provided a supportive home from which to conduct my research and see this project through to completion.

During my research I was fortunate enough to interview several members of the intelligence community and policy makers who shared the benefits of their wisdom and helped keep my research focused on matters that were relevant to praxis: James Bruce, Joseph Eash, Leon Fuerth, John Hanley, Barry Leven, Carmen Medina, Paul Pillar and Jennifer Sims are all deserving of more thanks than I can provide.

I would also like to thank the numerous scholars and practitioners whose time and correspondence, often quite extensive, enriched my research, provided an opportunity for discussing ideas, and offered important refinements to my thinking. Likewise, my peers and colleagues at Booz Allen Hamilton, National Defense University, BAE Systems, LMI, George Mason University, Icosystem, the Santa Fe Institute and elsewhere were all exceptionally generous. Amongst them, I owe David Cedel and Maciej Latek a special debt of gratitude for their technical assistance and support, while Elisa Bienenstock, Adam Korobow, Joseph Mait, Harold Scott and Wayne Zandbergen provided just enough prodding to keep me focused on finishing. I would also like to thank Karen Underwood and Colleen Spears for their assistance in getting me across the finish line.

The ideas of this project were planted many years ago and cultivated by four mentors who consistently pushed me to think about intelligence analysis and opportunities to improve its practice. My father, Forrest Frank, Art Hulnick, Barry Leven, and Desmond Saunders-Newton each provided encouragement and insights that allowed me to look past the limitations of my own experience and studies by reaping the benefits of theirs.

Finally, I need to thank my wife, mother, and grandmother for their endless support and creation of an environment that allowed me to pursue my studies for as long as I needed, and probably longer. Amy Frank, Ilene Friedman, and Bea Friedman made this possible.

Table of Contents

	Page
List of Tables	vii
List of Figures	viii
Abstract	x
Chapter 1 : Introduction	1
What is Agent-Based Modeling?	1
Agent-Based Modeling and Intelligence Analysis	3
Mapping the Argument	5
Chapter 2 : Models, Simulation, and Agent-Based Modeling	17
What are Models?	17
Models, Simulations, and Games	25
Bounded Rationality	30
The Limits of Mental Models	33
Types of Formal Models	40
The Uses of Models	57
Agent-Based Modeling	78
The Benefits of Agent-Based Modeling	88
Chapter 3 : Intelligence Analysis and Tradecraft	90
The Three-Level Problem of Intelligence Analysis	91
Producer/Consumer Relations	94
Structured Analytic Techniques and Analytic Tradecraft	130
Structured Analytic Techniques and the Manual Method	161
The Challenge of Strategic Intelligence	185
From Structured Analytic Techniques to Model-Centric Tradecraft	195
Chapter 4 : Intelligence and the Philosophy of Science	199
Intelligence and the Philosophy of Science	200
What is Science?	221
Science, Time, Induction, and Intelligence	256
Intelligence Challenges to Science	268
A Microscience for Intelligence Analysis	277
Chapter 5 : The Schelling Segregation Model	284
Thomas Schelling's Segregation Model	284
Schelling Segregation Model as an Analytic Tool	294
Extending the Core Segregation Model	312
Chapter 6 : Developing Theories and Hypotheses	314
Coercion and Community Formation	314

Enforcer Agents	315
Out-Group Experiments.....	321
In-Group Enforcement.....	329
Summary.....	338
Chapter 7 : Coordination between Collection and Analysis.....	343
Agent-Based Modeling and Intelligence Collection.....	345
Collection Capabilities.....	347
Collection Demonstration.....	354
Summary.....	387
Chapter 8 : Structure and Agency.....	389
Modifications for Microscience Experimentation	389
Experimenting with Agency and Structure.....	393
Analysis of Simulation Results.....	398
Summary.....	417
Chapter 9 : Conclusions	421
The Policy Pull.....	421
The Problem of Social Identity.....	429
Developing a Model-Centric Analytic Tradecraft	434
Future Research	438
Appendix.....	442
Interview with Leon Fuerth, September 21, 2011	442
Interview with Paul Pillar, February 1, 2012.....	449
Interview with John Hanley, February 9 2012	457
Interview with Carmen Medina, June 5 and July 3, 2012	464
Interview with James Bruce, August 13, 2012	468
Interview with Jennifer Sims, August 20, 2012.....	474
Interview with Joseph Eash, September 6, 2012	480
Interview with Barry Leven, October 25, 2012.....	486
References.....	500

List of Tables

Table	Page
3-1: Original Alternative Competing Hypotheses Table.	179
3-2: Contemporary Alternative Competing Hypotheses Table.	181
6-1: Out-Group Enforcer Parameters.	322
6-2: Total Time Statistics for Out-Group Enforcement Experiments.	325
6-3: Percentage Mixed Statistics for Out-Group Enforcement Experiments.	328
6-4: In-Group Enforcer Parameters.	330
6-5: Total Time Statistics for In-Group Enforcement Experiments.	332
6-6: Percentage Mixed Statistics for In-Group Enforcement Experiments.	337
7-1: Experimental Design for Sensor Simulation Demonstrations.	360
8-1: Experimental Parameters for Normal Movement Experiments.	395
8-2: Experimental Parameters for Purposeful Movement Experiments.	396
8-3: Statistical Measures for the Run Distance Metric for Each Free Agent (Normal Movement).	405
8-4: Statistical Measures for the Percentage Difference Metric for Each Free Agent (Normal Movement).	408
8-5: Statistical Measures for the Run Distance Metric for Each Free Agent (Purposeful Movement).	414
8-6: Statistical Measures for the Percentage Difference Metric for Each Free Agent (Purposeful Movement).	417

List of Figures

Figure	Page
2-1: Physical Models.....	41
2-2: Two Versions of the Intelligence Cycle	56
2-3: Emergence of Spatial Clusters in the Edge City Model	86
3-1: Sherman Kent's Intelligence Pyramid	126
4-1: Measurement and Estimates of Physical Constants over Time.....	211
4-2: Agency and Structure in Counterfactuals.....	274
5-1: Baghdad's Transition from Mixed to Segregated Communities.....	285
5-2: The Moore Neighborhood.....	289
5-3: Segregation Model Step	290
5-4: Segregation Model Initial and Final States	292
5-5: Segregation Model Response to Changing Population Density.....	293
5-6: The Interface of the Core Segregation Model	298
6-1: Depiction of Enforcer Agents.....	319
6-2: Total Time Steps for Out-Group Enforcement Simulations.....	323
6-3: Final Percentage Mixed for Out-Group Enforcement Simulations.....	327
6-4: Total Time Steps for In-Group Enforcement Simulations	331
6-5: Final Percentage Mixed for In-Group Enforcement Simulations.....	334
6-6: Sample Final Settlement Results of In-group Enforcement.....	335
7-1: Initial and Final State of the Simulation.....	355
7-2: Comparisons between Ideal Surveys and the True State of an Agent Population ...	356
7-3: Comparisons between Ideal Human Agents and the True State of an Agent Population.	357
7-4: Comparison Between a Perfect Technical Sensor and the Actual State of the Population	358
7-5: The Actual and Observed Percentage of Agents in Mixed Neighborhoods.....	363
7-6: Actual and Observed Percentage of Agents Happy with their Current Neighborhood for the Surveys and Human Agents	367
7-7: A Depiction of a Single Sensor in Tag Form	371
7-8: Panels Depicting the Status of a Genetic Algorithm for Developing Sensor Portfolios.....	379
7-9: The Percentage Happy and Percentage Mixed Metrics as Estimated by Sensor Portfolios for Generations 1, 15, and 30	380
7-10: Progression of the Population of Sensor Portfolios.....	382
7-11: Scatter Plots Depicting the Portfolio Cost and Fitness for Multiple Generations ...	384
7-12: Sensor Counts by Type for the Population of Portfolios by Generation.....	385

8-1: Initial Condition for the Segregation Model in All Experiments	399
8-2: Initial Positions of the Free Agents for All Experiments	400
8-3: Initial and Final State of Run 0 for all Normal Movement Experiments	401
8-4: Composite Results of Experiments 1.0-1.6	402
8-5: Composite Results of Experiment 1.5 and 1.6	403
8-6: The Distribution of the Run Distance Metric for Experiments 1.0-1.6.....	404
8-7: The Distribution of the Percentage Difference Metric for 10,000 Runs of Each Free Agent for Experiments 1.0-1.6.....	407
8-8: The Initial State and Run 0 that Established the Baseline for all Purposeful Movement Experiments.....	409
8-9: Composite Results of Experiments 2.0-2.6	411
8-10: Comparison between the Composite Results of Experiments 2.2 and 2.5	412
8-11: Distribution of the Run Difference Metric for Experiments 2.0-2.6.....	413
8-12: Comparison of Run Distance Measure for Free Agents 0 and 5	415
8-13: The Distribution of the Percentage Difference Metric for 10,000 Runs of Each Free Agent for Experiments 2.0-2.6.....	416

Abstract

AGENT-BASED MODELING AND INTELLIGENCE

Aaron B. Frank, Ph.D.

George Mason University, 2012

Dissertation Director: Dr. Robert Axtell

The United States Intelligence Community (IC) was born out of the experiences and organization of the Office of Strategic Services during World War II and became a permanent fixture of the national security establishment with the passage of the National Security Act of 1947.ⁱ Since its inception, there has been a strong fascination with the secret aspects of its work, particularly with respect to the clandestine collection of information and covert efforts to influence foreign governments, and to undermine rival intelligence services.ⁱⁱ By comparison, intelligence analysis, specifically the ways in which intelligence professionals develop and present assessments about the international

ⁱ Roger Z. George, "Central Intelligence Agency: The President's Own," in Roger Z. George and Harvey Rishikof, eds., *The National Security Enterprise: Navigating the Labyrinth* (Washington, DC: Georgetown University Press, 2011), pp. 157-158.

ⁱⁱ G. J. A. O'Toole, *Honorable Treachery: A History of U.S. Intelligence, Espionage, and Covert Action from the American Revolution to the CIA* (New York, NY: Atlantic Monthly Press, 1993); and Christopher Andrew, *For the President's Eyes Only: Secret Intelligence and the American Presidency from Washington to Bush* (New York, NY: Harper Perennial, 1996).

system to policy makers, has been relatively ignored.ⁱⁱⁱ As a result, intelligence analysis has remained largely under-theorized within the study of international relations, despite its prominent role in strategic thinking—only receiving significant attention in the aftermath of perceived failures.^{iv}

Understanding intelligence analysis is made more complex as new technologies affect analysts' access to new sources and methods. The challenge of technological change in intelligence studies has often been approached in one of three ways. New technologies may be seen as the source of new threats and opportunities in the international system, thus falling within the general application of security studies or international relations through threat assessments and estimates.^v A related, secondary study of technological change is more confined to the IC, pertaining to how new technologies might be exploited for collecting information, deceiving adversaries, and processing data.^{vi} A third consideration of technology within the IC is largely concerned

ⁱⁱⁱ Paul R. Pillar, *Intelligence and U.S. Foreign Policy: Iraq, 9/11, and Misguided Reform* (New York, NY: Columbia University Press, 2011); and Joshua Rovner, *Fixing the Facts: National Security and the Politics of Intelligence* (Ithaca, NY: Cornell University Press, 2011).

^{iv} Richard A. Posner, *Uncertain Shield: The U.S. Intelligence in the Throes of Reform* (Lanham, MD: Rowman & Littlefield, 2006); Richard K. Betts, *Enemies of Intelligence: Strategic Intelligence & American National Security* (New York, NY: Columbia University Press, 2007); and Robert Jervis, *Why Intelligence Fails: Lessons from the Iranian Revolution and the Iraq War* (Ithaca, NY: Cornell University Press, 2010).

^v Ernst R. May, ed., *Knowing One's Enemies: Intelligence Assessment Before the Two World Wars* (Princeton, NJ: Princeton University Press, 1986); Williamson Murray and Allan R. Millett, eds., *Calculations: Net Assessment and the Coming of World War II* (New York, NY: The Free Press, 1992); and Thomas G. Mahnken, *Uncovering Ways of War: U.S. Intelligence and Foreign Military Innovation, 1918-1941* (Ithaca, NY: Cornell University Press, 2002).

^{vi} John F. Kries, ed., *Piercing the Fog: Intelligence and Army Air Forces Operations in World War II* (Washington, DC: Air Force History and Museums Program, 1995); and Wayne Michael Hall and Gary Citrenbaum, *Intelligence Analysis: How to Think in Complex Environments* (Santa Barbara, CA: Praeger Security International, 2010).

with organizational structure and processes, particularly as they relate to the sharing and securing of information within and between the communities.^{vii}

This study considers a fourth perspective of technological change within the IC by examining how new technologies affect the tradecraft of analysis itself. In doing so, several perspectives are considered in order to understand how Agent-Based Modeling (ABM) can transform the dominant analytic paradigms that guide how intelligence analysis is currently practiced. While the majority of intelligence studies are empirical in nature, this project develops and advances a normative theory of intelligence analysis that argues for the careful consideration, development, and implementation of a model-centric analytic tradecraft, particularly for the production of strategic intelligence.^{viii} Rather than focus on prior cases of intelligence failures in order to demonstrate how intelligence analysts, collectors, managers, and policy makers have erred in the past, this project provides theoretical arguments in favor of establishing a new analytic tradecraft and demonstrates, if only in a rudimentary fashion, what this new approach may look like in practice.

In this study, several persistent debates within intelligence studies, familiar to a small group of scholars and theoretically minded professionals, are reimagined in light of the challenges and opportunities that new modeling and simulation capabilities provide. At its simplest, most conservative application, ABM extends the ongoing development of Structured Analytic Techniques (SATs), the centerpiece of contemporary analytic

^{vii} Thomas Fingar, "Office of the Director of National Intelligence: Promising Start Despite Ambiguity, Ambivalence, and Animosity," in Roger Z. George and Harvey Rishikof, eds., *The National Security Enterprise: Navigating the Labyrinth* (Washington, DC: Georgetown University Press, 2011), pp. 139-155.

^{viii} Richard K. Betts, *Enemies of Intelligence: Strategic Intelligence & American National Security* (New York, NY: Columbia University Press, 2007).

tradecraft. At the maximum, ABMs can reframe analytic practice, synthesizing the longstanding and opposing views as to whether intelligence analysis is an art or a science, and in the process transform the most challenging and interesting aspect of the intelligence profession—the relationship between intelligence producers and consumers.

Chapter 1: Introduction

Agent-Based Modeling (ABM) provides an ability to transform the practice of intelligence analysis. By improving the quality and clarity of analytic products and promoting greater transparency in the relationship between intelligence producers and consumers, ABMs offer benefits to intelligence analysts, managers, and policy makers that are not available from other methods. Given the opportunities it affords, ABM should be viewed as the centerpiece of a model-centric analytic tradecraft whose development would continue the growth of contemporary Structured Analytic Techniques (SATs) while incorporating fundamental changes in scientific practices that have emerged from the study of complex adaptive systems, specifically, and non-linear, complex systems more generally.

What is Agent-Based Modeling?

ABM is a type of computational simulation used to study the properties of complex adaptive systems. Its development marks the convergence of several different lines of research that represent a diverse array of disciplines including computer science,

physics, chemistry, biology, ecology, economics, sociology, and psychology.⁹ By employing computational algorithms to represent systems as collections of interacting units, each capable of possessing unique attributes, behaviors, and experiences, the dynamics and properties of complex adaptive systems have become accessible to researchers. Within the scientific community, this has facilitated interdisciplinary research that crosses multiple problem domains, linking micro and macro levels of systems where emergent properties, phase transitions, path dependence, feedback, hysteresis, and other dynamics occur.

Within the social sciences, the scholarly fields most closely related to intelligence analysis, ABMs have been employed to address the linkages between individual decision-making and collective outcomes on matters as diverse as patterns of human settlement,

⁹ Per Bak, *How Nature Works: The Science of Self-Organized Criticality* (New York, NY: Copernicus, 1996); Mark d'Inverno and Michael Luck, *Understanding Agent Systems* (New York, NY: Springer, 1998); Simon Levin, *Fragile Dominion: Complexity and the Commons* (Cambridge, MA: Perseus Books, 1999); Gerhard Weiss, ed., *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence* (Cambridge, MA: MIT Press, 1999); Timothy A. Kohler and George J. Gumerman, eds., *Dynamics in Human and Primate Societies: Agent-Based Modeling of Social and Spatial Processes* (New York, NY: Oxford University Press, 2000); Scott Camazine, Jean-Louis Deneubourg, Nigel R. Franks, James Sneyd, Guy Theraulaz and Eric Bonabeau, *Self-Organization in Biological Systems* (Princeton, NJ: Princeton University Press, 2001); Nigel Gilbert and Rosaria Conte, eds., *Artificial Societies: The Computer Simulation of Social Life* (New York, NY: Routledge, 2003); Andrew Ilachinski, *Artificial War: Multiagent-Based Simulation of Combat* (New Jersey, NJ: World Scientific, 2004); Volker Grimm and Steven F. Railsback, *Individual-Based Modeling and Ecology* (Princeton, NJ: Princeton University Press, 2005); Joshua M. Epstein, *Generative Social Science: Studies in Agent-Based Computational Modeling* (Princeton, NJ: Princeton University Press, 2006); Thomas Salamon, *Design of Agent-Based Models* (Czech Republic, Tomas Bruckner, 2011); and Flaminio Squazzoni, *Agent-Based Computational Sociology* (West Sussex, United Kingdom: James Wiley & Sons, 2012).

voting behavior, state formation and nationalism, military combat, finance, and more.¹⁰

By applying these methods to strategic intelligence problems, analysts can reap many benefits from formal models that simulate the interactions of adaptive, strategic actors in order to identify potential futures of interest to policymakers, while simultaneously improving the relationships between intelligence producers and consumers that ultimately determine the value of intelligence.

Agent-Based Modeling and Intelligence Analysis

ABM is distinctive because of the depth and breadth of its potential to transform the practice of intelligence analysis. Whereas most research methods find application to niche areas within the intelligence community, the prospective uses of ABM in intelligence are virtually unlimited. The ability to represent intelligence targets as complex adaptive systems provides high levels of representational flexibility that allow ABMs to represent the mental models of analysts and policy makers. Likewise, because ABMs are simulations, they enable the systematic generation and exploration of alternative futures and scenarios, making them epistemologically compatible with the ways in which policy makers identify and select between alternative options.

At the analytic level, ABM provides new opportunities to study the behavior of complex adaptive systems composed of interacting, learning actors without making the

¹⁰ Lars-Erik Cederman, *Emergent Actors in World Politics: How States and Nations Develop & Dissolve* (Princeton, NJ: Princeton University Press, 1997); Timothy A. Kohler and George J. Gumerman, eds., *Dynamics in Human and Primate Societies: Agent-Based Modeling of Social and Spatial Processes* (New York, NY: Oxford University Press, 2000); Andrew Ilachinski, *Artificial War: Multiagent-Based Simulation of Combat* (River Edge, NJ: World Scientific, 2004); Vincent Darley and Alexander V. Outkin, *A NASDAQ Market Simulation: Insights on a Major Market From the Science of Complex Adaptive Systems* (Hackensack, NJ: World Scientific, 2007); and Michael Laver and Ernest Sergenti, *Party Competition: An Agent-Based Model* (Princeton, NJ: Princeton University Press, 2012).

“heroic assumptions” that have limited the adoption of other modeling formalisms.¹¹

This underlying approach to framing problems aids analysts in the same way as contemporary Structured Analytic Techniques (SATs) by encouraging the explication of assumptions and articulation of relations between actors observed in data or theorized by analysts. However, because ABMs also enable the simulation of these relations, systems whose behaviors are too complex to be examined without cognitive assistance can be rigorously explored. This methodological capability encourages analysts to augment empirical assessments with computational simulations in order to generate alternative futures and identify potential threats and opportunities.

The improved ability to study complex adaptive systems provided by ABMs qualifies it as an important methodology that should be integrated into analysts’ set of research tools. However, this does not constitute a transformative capability in its own right. Indeed, analysts already have access to a variety of methods, both quantitative and qualitative, that have proven immensely important for particular problems, but have not dramatically transformed analytic tradecraft itself. This is because the institutional, cultural, and epistemological circumstances of strategic decision making create a context that strains relations between intelligence producers and policy-making consumers.

It is at the boundary of intelligence analysis, specifically the interface with policy makers, that the transformative potential of ABM will ultimately be determined. Because the underlying ontological structure of ABMs matches the agent-oriented worldviews of decision makers, analytic tradecraft that develops and examines simulations of their

¹¹ Herbert A. Simon, *The Sciences of the Artificial* (Cambridge, MA: MIT Press, 1996), pp. 25-49.

worldviews can facilitate an extended dialog between producers and consumers that creates the necessary foundation of trust and personal relationships that undergird the effective use of intelligence in policy and strategy. Thus, ABMs are not only tools for scientific and strategic research, but also mechanisms for facilitating dialogs across the intelligence community's many institutional barriers.

The full realization of ABMs' transformative potential requires recognizing modeling and simulation as an integral part of the analytic tradecraft—operating at the center of research and production while touching the boundaries of organizational processes including coordination across analytic units and organizations, working with collectors, and developing and managing relations with consumers. This means viewing ABM as more than a particular analytic technique whose use is framed in strictly methodological terms. Instead, it requires that ABM be seen as the centerpiece of a model-centric tradecraft of intelligence analysis that gives equal emphasis to both the substance of available intelligence and process of modeling and simulation. Together, ABMs can inform particular analytic judgments by generating synthetic data that can augment the empirical evidence undergirding analysts' inferences, while the process of developing models and exploring their behavior can improve consumers' confidence in the rigor, neutrality, and transparency of intelligence products.

Mapping the Argument

This study evaluates ABM within the context of intelligence, international relations, and national security policy making. Its primary purpose is integrative—merging concepts and insights from several disparate fields into a single, coherent

approach to intelligence analysis enabled by computation. Therefore, each of the following chapters should be seen as part of a larger progression toward the conclusion that ABM can provide transformative analytic capabilities for the intelligence community, but that the realization of its potential requires focusing on ABM's scientific and analytic merits with respect to the study of complex adaptive systems, and its fit within the institutional structure of the intelligence community.

Chapters 2 through 4 provide the conceptual foundation for the development of a model-centric analytic tradecraft developed around the employment of ABM. These chapters each address different aspects of the argument by examining a combination of the historical record of intelligence successes and failures, methodology, and philosophy. Afterward, Chapters 5 through 8 provide demonstrations of several of the points raised in prior chapters, particularly the role of models in theory development, the detection of collection biases, and the boundary between agency and structure in complex systems. Finally, Chapter 9 concludes the project by discussing why the capabilities provided by ABM are important for contemporary intelligence problems, and important issues that must be resolved in future research.

Chapter 2 provides an introduction to modeling and simulation in general, and ABM more specifically. It introduces many of the central concepts associated with employing models as part of analytic tradecraft within the intelligence community, including many of the ways models are employed as tools for research, analysis, and other organizational activities. Moreover, it highlights how ABM differs significantly from other approaches to formal research methods, most notably mathematical models.

These differences stem from the way in which actors are represented in each modeling formalism, the most important of which concerns the ways in which units in a system are aggregated in the modeling process. Whereas mathematical models compress populations of actors into a singular representative entity, usually based on the group's mean, ABM allows for each actor to remain discreet and exist as separate units within the model. Because each actor is individually represented, new opportunities exist for considering how their heterogeneous properties, whether in the form of alternative attributes, behaviors, or individual experiences, affect the system as a whole.

The emphasis on disaggregation and individual representation creates new research opportunities for analysts. By shedding methodologically imposed constraints of rationality and homogeneity the defining features associated with particular cases or questions can be captured by ABM. This means that salient features of structure and agency can be investigated via simulation, establishing links between the micro and macro levels of analysis. By linking perspectives on individual decision making and action with structural and systemic results, the essential elements of strategic decision making and interests of policy makers can be satisfied. By identifying relationships between choices and outcomes, intelligence producers and consumers can improve their search for potential threats and opportunities—uncovering actions that might mitigate the damage of adversarial actions while improving the ability to shape the international system in desired ways.

Chapter 3 continues the discussion of models and modeling by considering the institutional, cultural, and substantive context of intelligence analysis. While the

intelligence community has largely eschewed formal modeling in the production of strategic intelligence, every study of intelligence failures and successes over the last few decades has emphasized the importance of analytic methodology. This has motivated the development and increasing use of SATs, which are now a centerpiece of analytic tradecraft. What has gone largely unrecognized has been the fact that SATs are in fact a particular type of modeling, consistent with the processes involved in designing, developing, and examining ABMs.

The evolution of analytic tradecraft has emphasized the employment of SATs for two reasons: improving the rigor and consistency of analytic reasoning, and increasing the transparency of analysts' assumptions, use of evidence, and inferences in order to make intelligence products more acceptable to consumers. In both instances, ABM continues the logical progression of tradecraft's development.

The primary goal of SATs is providing analysts with tools to cope with bounded rationality, i.e. the inherent cognitive limits of the human brain. This is accomplished by challenging the natural tendency to satisfice, and to filter assumptions and interpretations of information based upon how well they confirm existing beliefs and expectations about the international system. The techniques are overwhelmingly manual, light-weight, problem agnostic approaches to articulating assumptions, organizing data, and communicating the relationships between evidence and inference in order to make the analytic judgments and their justification transparent to policy makers.

ABMs extend the application of SATs by addressing two of their important shortcomings. First, because ABMs are simulations that enable the examination of

complex adaptive systems, they allow for the static conjectures of analysts to be animated, creating dynamic representations of intelligence targets that can be studied *in silico*. This furthers the defense against bounded rationality by employing computation to improve the performance of mental simulations by highlighting and mitigating errors that result from the brain's cognitive limits of coping with complexity. Secondly, ABMs allow for the incorporation of domain-specific knowledge into specifications of agents and their respective decision-making and behavioral rules. This encourages the introduction of social science theories to augment the mental models of analysts and policy makers in analysis. Therefore, ABMs provide a platform for drawing upon political science, economics, sociology, psychology, anthropology, etc., adding substance to the generic problem-solving heuristics provided by SATs.

Incorporating ABM into analytic tradecraft also requires analysts to develop and employ new research skills. Existing tradecraft emphasizes the mastery over the details of particular problems, such as a country's geography, culture, and history. Relying on modeling and simulation in the production of intelligence introduces a need for abstracting from reality, developing potential models, deriving implications from model assumptions, and comparing the operations of abstract systems with empirical reality and future collection.¹²

Chapter 4 draws upon the philosophy of science in order to show why ABM is particularly valuable for intelligence tradecraft. The intelligence community has long debated whether analysis is an art or science, with most practitioners and scholars

¹² James G. March, "Making Artists Out of Pedants," in Ralph M. Stogdill, ed., *The Process of Modeling-Building in the Social Sciences* (New York, NY: W.W. Norton & Company, 1970), pp. 63-65.

believing in the former rather than the latter. However, the justifications for this conclusion rest on a flawed definition of science and understanding of how it is practiced. By examining several different demarcation criteria found in the philosophy of science, many of the features of intelligence analysis that are assumed to distinguish it from science are shown to be consistent across both domains, calling into question the art vs. science dichotomy.

The similarities between science and intelligence are not absolute, and in two largely unacknowledged ways, more different than recognized. While philosophers have had many important disagreements regarding what science is and how it is practiced, they generally agree on many important issues that fundamentally challenge the practicality of intelligence analysis and the informational needs of policy makers by calling into question the limits of *a priori* knowledge. By accepting the primacy of inductive inference, empiricism, and generalization, the epistemological justifications of scientific knowledge rest on *a posteriori* experience, where data from experiments and observations serve as the ultimate arbiters of claims to knowledge.

In contrast, intelligence analysts and policy makers must look into the future and consider a range of alternative outcomes that might result from how their choices interact with those of other actors in the international system. Rather than search for and explain generalizable patterns across a universe of cases, as science emphasizes, intelligence producers and consumers seek to identify what futures might be theoretically possible given a particular case or set of circumstances. As a result, the fundamental orientation of their research and analysis must emphasize agency over structure in order to unpack

how actors' choices interact and allow for differentiating between those outcomes that are structurally determined and those that are contingent on agents' strategic choices and actions.

ABM is well suited for meeting the needs of intelligence analysts and policy makers. Its ability to represent the features and actors in particular cases can serve as the basis for generating alternative scenarios or examining counterfactuals in order to identify causal paths and sources of influence within complex adaptive systems. This allows for the development and practice of a microscience, an experimental, model-based approach to research and analysis that adds methodological rigor to the development, exploration, and assessment of the relationship between structure and agency.

The development of an ABM-enabled microscience helps close the gap between the practice of intelligence analysis as both art and science. It facilitates the systemic examination of the international system and other intelligence targets by encouraging the modeling and simulation of particular cases in order to identify the boundaries over which policy makers' choices can shape outcomes. By emphasizing the study of particular cases rather than populations of them, the analytic balance is tilted toward specificity and employing knowledge that is already the professional currency of the analytic community. The use of computational simulation to generate new scenarios and explore counterfactuals provides analysts with an ability to generate information that cannot be collected in the real world. This improves the justification of their assessments by transforming the opaque practice of mental simulation into a transparent study of artificial systems used to develop and test hypotheses.

Chapter 5 transitions from theory to application, demonstrating how the concepts discussed in the previous chapters can be operationalized. It provides a description of what is generally considered to be the first ABM, Thomas Schelling's model of racial segregation. This simple model is employed as the basis for articulating how ABMs are designed, implemented, and manipulated. The employment of Schelling's segregation model is used because of its simplicity when compared with other more sophisticated ABMs that have been used in scientific research and other analytic endeavors. By using the segregation model as a pedagogical device, several of the different ways ABM can be employed in intelligence tradecraft are shown in the subsequent chapters.

The discussion of the segregation model provides more than a demonstration of the technical operations of ABMs. It also shows how the paradigm implicit in intelligence analysis, and international relations theory more broadly, is not sufficient to address many real-world intelligence and policy problems. Whereas the leading analytic paradigm emphasizes the evaluation of a target's capabilities and intentions, the challenge of complex adaptive systems is that the processes by which capabilities and intentions are aggregated or interact may be equally or more important. The demand for a new analytic paradigm raises the prospect of new collection targets and analytic considerations in order to justify intelligence assessments. Chapters 6-8 provide demonstrations of ABM's use in analytic tradecraft by extending the segregation model in different ways.

In Chapter 6, the most typical employments are demonstrated by using simulations to generate new theoretical insights that can inform how analysts frame

intelligence problems, relate hypotheses to the available evidence, and consider what information they should request from collectors. The specific modifications introduced in Chapter 6 involve the consideration of coercive force within society. Schelling's original model operated under the assumption that agents' behaviors were entirely voluntary and therefore accurately reflected their individual preferences. However, in many areas where sectarian or ethnic conflict exists, violence or the threat of violence is often employed.

The addition of coercive power to the segregation model provides a demonstration of how ABMs can be modified by analysts to investigate particular questions. In the case of coercion, this investigation assists in the process of hypothesis generation and testing by suggesting new means for relating evidence and inference regarding group solidarity, strength, and the use of force, as well as introducing new theoretical concepts, such as the aggressiveness of coercive actors, that are realized based on making technical changes to the model.

Chapter 7 examines how ABM can bridge gaps between intelligence analysts and collectors. While numerous scholars and practitioners have noted the significance of having analysts and collectors work together, the relationship between the two is regarded as largely ad hoc and idiosyncratic. By simulating intelligence collection within ABM's artificial societies, analysts and collectors can prioritize targets, identify sources of collection bias, and design collection strategies tailored to the particular needs of analysts.

The simulation of intelligence collection is performed within the segregation model in order to demonstrate the operations of three different sources of information available to analysts. The use of surveys, human informants, and technical sensors each constitute different collection capabilities that reveal common and uncommon modes of biases with which analysts must familiarize themselves. Moreover, several computational strategies can be employed to search for novel collection capabilities or mixed strategies that fuse multiple sensors together in order to provide analysts with a perspective on a targeted system that may not be derived from a single source. By employing evolutionary computation in the design of collection strategies and capabilities, collection approaches that might be well suited for particular requirements could be identified via simulation that might have otherwise gone unrecognized.

Chapter 8 provides a demonstration of the power and current limitations of ABM with respect to the fundamental purpose of intelligence and policy analysis: uncovering the linkages between actions and outcomes in complex adaptive systems. In order to accomplish this goal, the segregation model is reimplemented to establish experimental control over individual agents in the system in order to investigate how differences in their own decision making affect the system as a whole.

This final demonstration identifies the forefront of ABM with respect to design and modeling in support of intelligence and policy. It shows how the full capabilities and potential of ABM remain an untapped resource. Thus, it also serves as a call to reevaluate how the modeling community currently employs ABM and designs experiments.

Finally, Chapter 9 summarizes the arguments presented in prior chapters, and examines how a model-centric analytic tradecraft might be applied to pressing national security and intelligence problems. This examination also notes how important the opportunities provided by ABM are to the intelligence community given the increasingly diverse array of subjects, sources of threats, and opportunities the community must address.

These final points regarding the importance and potential of ABM in intelligence analysis are shown by discussing the challenge of understanding identity in the international system. Identity is a complex theoretical concept and its study challenges the dominant theories and methods of international relations. Identity is constructed through social interaction and evolves over time, making it dynamic, adaptive, and not under the rational control of individuals. Because identity links individuals and groups together, it cannot be reduced to the study of individuals or groups in isolation, thus defeating conventional analytic treatments based on decomposing systems into independent units, studying them in isolation, and then recombining them into a whole. Finally, because identities include idealized models of behaviors that establish norms and heuristics, they limit the extent to which actors can be deemed rational in the classical sense. Thus, once agents are regarded as identifying with particular social, economic, political, or other groups, they cannot be regarded as perfectly rational actors and must be treated as boundedly rational and subject to the information, preferences, and standards of behavior of their salient groups and roles. For these reasons, ABMs provide analysts

with an opportunity to examine the dynamics of identity politics, and other problems, in ways that have proven difficult via other means.

In conclusion, ABM has the ability to transform analytic tradecraft. It continues the natural evolution of SATs, while serving as a vehicle for introducing new critical reasoning skills, theory, and structure into the generation and assessment of scenarios and counterfactuals. This provides new opportunities for analysts to systematically explore the consequences of alternative choices that adversaries and consumers might consider, and facilitate an extended dialog with policy makers who can share in the design and exploration of models' behaviors. ABM's transformation potential stems from its underlying ontological commitments to disaggregation, heterogeneity, and decision making, which allows for the examination of structure and agency in circumstances that lie beyond the limits of empirical analysis.

Chapter 2: Models, Simulation, and Agent-Based Modeling

This chapter examines the design and use of models as research and analytic tools. It provides a working definition of what models are, identifies different classes of models, and places modeling in the broader context of scientific inquiry. Also, it discusses many of the ways that models are employed by individuals and organizations, with a specific emphasis on their non-predictive uses. Finally, it provides an introduction to ABM, introducing the reader to the analytic concepts and model structures that provide transformational potential.

What are Models?

In their most general form models are representations of systems. While simple, this statement does not convey the full importance or diversity of roles played by models in scientific research, intelligence analysis, and human decision making more generally. Efforts to provide more precise definitions of models add more detail, but introduce sources of disagreement because they are imbued with normative considerations about how models should be developed, what models should accomplish, and how they should be used.

To examine the difficulties of determining what models are and uncover how they contribute to science, Daniela Bailer-Jones interviewed several prominent scientists and asked how they defined models. She received many different responses, each of which

contained a core set of commonalities as well as important, divergent indications of how they were used in research, including several respondents who noted the term model was itself too difficult to define.¹³ Those definitions that were offered varied considerably, ranging from describing models as the ways in which problems are characterized and set up in order investigate them, to formal statements made in the form of testable hypotheses, to summaries of a system's fundamental properties and assumptions that, together, allow for input data to predict output data. Some respondents argued that models should be mathematical, made of precise numerical or logical statements made available for investigation, while others argued that they might be paper diagrams or physical constructs that represented something other than what they were.¹⁴

The diversity and disagreements that appeared in Bailer-Jones's interviews were not unique. Almost any examination of how philosophers and scientists define the term model will largely replicate the diversity and inconsistency of Bailer-Jones's sample. In the social sciences, the scholarly disciplines that most closely resemble intelligence analysis, models are primarily seen as surrogate systems that facilitate the examination of real-world situations and phenomenology that are too complex or uncontrollable to study and control directly. As a result, models possess two essential features.

First, they are abstractions or simplifications of the real world. Charles Lave and James March highlighted the importance of simplicity of models as abstract, partial representations of systems or problem when they noted:

¹³ Daniela M. Bailer-Jones, *Scientific Models in Philosophy of Science* (Pittsburgh, PA: University of Pittsburgh Press, 2009), p. 6.

¹⁴ Daniela M. Bailer-Jones, *Scientific Models in Philosophy of Science* (Pittsburgh, PA: University of Pittsburgh Press, 2009), pp. 5-13.

A model is a simplified picture of a part of the real world. It has some of the characteristics of the real world, but not all of them. It is a set of interrelated guesses about the world. Like all pictures, a model is simpler than the phenomena it is supposed to represent or explain.¹⁵

Because models are simplifications of systems, they can be studied with greater ease, and often manipulated in order to explore behaviors and properties that could not be observed in the real world. Thus, the second critical feature of models is their ability to support inferences about target systems through the study of alternative or surrogate systems.

The treatment of models as a means for studying inaccessible or uncontrollable systems indirectly was noted by Jim Doran and Nigel Gilbert, who wrote that:

There are many and various definitions of modeling. The notion we shall use is a simple one. We wish to acquire knowledge about a target entity *T*. But *T* is not easy to study directly. So we proceed indirectly. Instead of *T* we study another entity *M*, the “model,” which is sufficiently similar to *T* that we are confident that some of what we learn about *M* will also be true of *T*.¹⁶

These two definitions are quite broad with respect to how models are constructed and employed. They offer no normative claims as to whether models should be mathematical equations, diagrams, computer programs, or other formalisms. Moreover, while they insist that models should allow for inferences to be made about the systems they represent, these definitions are agnostic as to whether those inferences must be predictions about the real world or insights arrived at by isolating particular features of systems that cannot be observed empirically.

¹⁵ Charles A. Lave and James G. March, *An Introduction to Models in the Social Sciences* (Lanham, MD: University Press of America, 1993), p. 3.

¹⁶ Jim Doran and Nigel Gilbert, “Simulating Societies: An Introduction,” in Nigel Gilbert and Jim Doran, eds., *Simulating Societies: The Computer Simulation of Social Phenomena* (London, UK: University College of London Press, 1994), p. 4.

With respect to the social sciences and intelligence community, the predictive applications of models have proven problematic given the complexity of the systems and questions they seek to understand. However, predictive difficulties are not necessarily a significant problem in the context of policy making, at least when compared to their scientific uses, because the goal of intelligence analysts is to assist consumers in making better decisions, whereas scientists tend to compete over developing the best model, often evaluated by its predictive power. While science largely progresses through the development of increasingly improved models, in the context of intelligence production and consumption models are incidental and ancillary, i.e., they are tools that are developed and employed on the way to meeting the real goal of informing decision makers.

Abstraction, Simplification, and Intelligence Analysis

The key elements of modeling—abstraction and simplification, exploration, and derivation—demand a set of skills that challenge the traditional emphasis of intelligence analysis, and the social sciences against which analytic practices most closely resemble. The basic skills associated with intelligence analysis have often drawn on the social sciences, and generally emphasize reading, organization, and categorization as the hallmarks of accomplishment and expertise. Thus, the currency of intelligence analysts' expertise is their knowledge of the relevant literature, academic and intelligence reports,

and particular facts and history of problems, such as a country's geography, culture, and history.¹⁷

By comparison, modeling draws on a different skill set, which is often unrecognized and undervalued within the intelligence community because of the primacy of current intelligence—providing consumers with descriptions of ongoing events and activities of intelligence targets—over strategic intelligence, which provides consumers with an assessment of what is possible or could happen.¹⁸ As long as intelligence producers and consumers emphasize current intelligence, then the mastery over the specifics of intelligence problems will remain essential and assessed as the hallmark of analytic expertise. However, in the cases of strategic intelligence, where the character of analysis changes from describing current events to estimates about potential futures, analysis becomes increasingly theoretical and rests on the development, exploration, and implications of alternative conceptual frameworks through which intelligence information and problems are viewed. Under these circumstances, strategic intelligence resembles the process of model building and assessment, where the analysis depends on the ability to abstract from reality, develop potential models, derive implications from models, and compare the behaviors and outputs of models with information about target systems.¹⁹

Modeling rests on abstraction and simplification, activities which necessarily leave out features of specific problems in their representation of target systems. In the

¹⁷ James G. March, "Making Artists Out of Pedants," in Ralph M. Stogdill, ed., *The Process of Modeling-Building in the Social Sciences* (New York, NY: W.W. Norton & Company, 1970), p. 63.

¹⁸ The differences between current and strategic intelligence will be discussed in greater detail in Chapter 3.

¹⁹ James G. March, "Making Artists Out of Pedants," in Ralph M. Stogdill, ed., *The Process of Modeling-Building in the Social Sciences* (New York, NY: W.W. Norton & Company, 1970), pp. 64-65.

context of intelligence analysis, abstraction and simplification are problematic for cultural reasons because the analysts' expertise is demonstrated by mastery over the details and nuance of particular cases. For example, a model that examines a country's commercial environment based on economic and technological trends, but leaves out detailed information on patterns of elite ownership and political connections, would be seen as making inadequate use of analysts' knowledge.

Another challenge of abstraction and simplification in the intelligence community is operational. Intelligence data, particularly highly detailed, localized information from Human Intelligence (HUMINT) and Technical Intelligence (TECHINT) sources, may be collected with great effort, expense, and risk. Not employing all available intelligence because it does not fit within an abstract and theoretical model can be seen as failing to maximize the intelligence community's resources, potentially stoking conflicts between intelligence collectors, analysts, and consumers.

In an environment where intelligence may be scarce, analysts may resist discarding what little information they possess.²⁰ Indeed, the desire to use all available information, particularly when dealing with hard targets that actively seek to deny or manipulate intelligence collection, can actually increase the risk of deception. Thus, an irony of intelligence analysis is that in cases where particular pieces of information are deemed critically important, the importance and need for abstract models are elevated in

²⁰ Robert Jervis, "Improving the Intelligence Process: Informal Norms and Incentives," in Alfred C. Maurer, Marion D. Tunstall, and James M. Keagle, eds., *Intelligence: Policy and Process* (Boulder, CO: Westview Press, 1985), p. 121.

order to ensure that scarce information is examined from as many perspectives as possible.

The tension between intelligence expertise, which emphasizes detail and nuance, and model building and design, which emphasizes abstraction and simplification, results from a misunderstanding in what models are and how they can be employed in analytic tradecraft. If models are treated as independent objects that make objective statements about the systems they represent, then they will be seen to directly challenge analytic expertise and experience. This challenge of employing models will be even greater if the simplicity of their representation is assumed to stand for the simplicity of the target system they represent. As long as intelligence analysts and consumers view models as objective and truthful, two extremes will likely dominate intelligence production and use. Enthusiasts may mistake the availability of models as a means for replacing analysts, who possess motivated and unmotivated biases, with models that are believed to provide objective statements about reality, and accept their outputs uncritically. Skeptics may view models as threats to roles played by analysts, or even decision makers, and therefore fiercely resist their employment in analytic and decision-making processes, denying producers and consumers their potential benefits.

Models are simultaneously dependent and independent of their creators. When formalized, they exist outside of the minds of their creators, and are therefore artifacts with their own unique properties and powers of agency.²¹ However, models remain

²¹ Margaret Morrison, "Models as Autonomous Agents," in Mary S. Morgan and Margaret Morrison, eds., *Models as Mediators: Perspectives on Natural and Social Science* (New York, NY: Cambridge University Press, 1999), pp. 38-65.

collections of guesses, conjectures, or assumptions about the world, and are therefore products of their creators' minds and ideas, sitting atop a foundation of their designers' beliefs, experience, and expertise. Therefore, models are best viewed as the externalized expressions of their developers' beliefs, whether those beliefs are about the behavior of actors and relations in a system, or in the selection and coding of data, and not as objective representations about the real world.

Because models are simpler than the systems they represent, they allow for insights to be gained from observing and manipulating them in ways that the real world does not allow. However, because models are simplifications of more complex systems, they do not behave exactly as the real world system does. One of the most famous statements about models, attributed to the statistician George Box, succinctly summarized the challenge of developing and using models by noting that "All models are wrong, some are useful." Skillful research and analysis hinges on the ability to effectively use models by understanding their limitations and the ways in which their behavior is likely to deviate from real-world systems.

Abstraction and simplification create an important contradiction if models are viewed as sources of objective predictions. Because models do not contain the full complexity of the actual system under investigation, anything that has been left out of the model is diminished and interpreted as having a value of zero. Importantly, however, this is a misleading conclusion since the missing variables, actors, processes, and other features only have no value in the artificial world of the model, not the real one it represents. Thus, the important virtue of parsimony often sought in model design and

experimentation must first and foremost be viewed as a practical consideration rather than an epistemic truth about reality.²²

The practical nature of parsimony in models vs. reality is an essential one with respect to model use and intelligence. For modeling purposes, it is important to recognize that there are many ways to deal with highly complex problems; formal models are most effective when viewed as a source for partial answers, often due to their inability to generate particular outcomes, isolate the ways particular variables or relationships operate, or construct idealized systems that can suggest the importance of variables or relationships through their exclusion. Simple, parsimonious models may be essential to the analytic process, but their presence should not be mistaken for believing that the target system itself is simple.

Models, Simulations, and Games

Understanding that models are representations of target systems, questions emerge as to whether they contain moving parts, change their properties over time, or possess interacting subsystems that produce emergent properties. Models that change over time are referred to as simulations. Simulations are dynamic, and show how actions and interactions within a system can change the system's structure and behavior over time. Alternatively, games are models that are social in nature, where outcomes are contingent on the interdependent choices of the system's actors. Games may be single-shot models, such as the classic prisoner's dilemma in game theory, or dynamic

²² John Lewis Gaddis, *The Landscape of History: How Historians Map the Past* (New York, NY: Oxford University Press, 2002), pp. 62-66.

simulations that involve competition, cooperation, and coordination between actors. In each case, when actors are interdependent, the system is strategic—meaning that each actor must consider what others might do when making their own decisions, resulting in the need for communication and deception.²³

The differences between models, simulations, and games can be subtle, and often rest on how a model is used in the analytic process. For example, the planning of the military raid that killed Osama Bin Laden (OBL) provided examples of the differences between these three terms. Intelligence analysts had constructed several different models of Bin Laden’s Abbottabad compound, including a computer model, a physical model that fit on a tabletop for military planners to analyze and use for illustrations in briefings, and a full-scale replica of the property and its buildings to aid in mission rehearsals.

These three models were described by Peter Bergen:

...the National Geospatial-Intelligence Agency used its minute reconnaissance of the compound to produce a digital computer-aided design (CAD) file of the type that engineers use in drawing up blueprints. From that CAD file, a four-by-four-foot model of the suspected bin Laden compound was constructed that was accurate down to the last tree. The model even included two tiny toy cars that represented the white Jeep with a spare tire and the red van that the Kuwaiti and his brother drove. The model became a vital prop at the CIA and White House for discussions about who was living on the compound and in which location, and later for talking through the planning of the various military options. General James Cartwright, then the vice chairman of the Joint Chiefs, recalls, “That was a good vehicle for us as we planned the various options to then sit down with that model and say... ‘This is how we would come at it; this is what would happen in this courtyard or this house.... Here’s how we would have more than one avenue of approach on what we thought were the target-inhabited buildings....’”

While the White House continued to debate the various COAs [Courses of Action] over the course of five days in early April, the SEAL

²³ Thomas C. Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1980).

team from DevGru's Red Squadron began its rehearsals on full-scale models of the compound in a secret facility deep in the forests of North Carolina. They practiced on a one-acre replica of the Abbottabad compound, fast-roping down from Black Hawks onto the courtyard of the compound and the roof of its main building. These rehearsals were observed by the overall commander of Special Operations, Admiral Eric Olson, a thoughtful Arabic speaker and former Navy SEAL, and by Mike Vickers from the Pentagon, Admiral McRaven, and Jeremy Bash of the CIA. The rehearsals took place in daytime and didn't include a practice run of the helicopter ride into Abbottabad, focusing only on what the SEAL team would do "on target...."

The SEAL teams rehearsed again for a week in mid-April, in the high deserts in Nevada, which replicated the likely heat conditions and the elevation of Abbottabad, which sits at four thousand feet. This time they rehearsed the entire mission from nighttime takeoff to the return to base more than three hours later. Again, Olson, McRaven, Vickers, and Bash observed the rehearsal, this time joined by Admiral Mullen. The observers were taken into a hangar, where the SEALs walked them through a "rehearsal of concept" drill using a cardboard model of the compound. The SEAL teams then flew off in their helicopters for about an hour. When they returned, the outside observers, now wearing night vision goggles, watched them as they assaulted the compound. During this rehearsal, wind conditions forced the helicopters to arrive at the target from an unexpected direction. This reminded the observers that no matter how many times the assault was rehearsed, there were still going to be some "game-time" decisions to make. The rehearsals also showed that the whole operation on the ground could be conducted in under thirty minutes—the amount of time the Pentagon had determined that the SEALs would have before they were interrupted by the arrival of Pakistani security forces.²⁴

These models of the compound reveal the ways model use can often determine the differences between models, simulations, and games. The reconstructions of the compound in electronic and physical form were static models. However, these static models became dynamic simulations when used by military planners as they simulated how Special Forces operators might approach the facilities, and infiltrate and exfiltrate

²⁴ Peter L. Bergen, *Manhunt: The Ten-Year Search for Bin Laden: From 9/11 to Abbottabad* (New York, NY: Crown, 2012), pp. 164-165, and 182-184.

under nighttime conditions. Then, by considering how best to complete the operation given potential interference from Pakistani security forces and the active resistance from the compound's inhabitants, these simulations became games requiring strategic decision making regarding the potential actions of competitors.

As the planning of the bin Laden raid showed, models, simulations, and games have been staples of military planning and defense analysis for decades. Indeed, from the very invention of the computer, the simulations of particular weapon systems and their uses have been a major source of research and investment for the Department of Defense (DOD).²⁵ Although the gap may be considered large between traditional wargames played by maneuvering small pieces on game boards in a fashion similar to chess, or via field maneuvers and exercises by military units, and computer models that employ complex mathematical or algorithmic logic, they reveal the diversity of ways in which modeling, simulation, and gaming can be operationalized, often in pursuit of the same insights.

Despite the importance of distinguishing between models, simulations, and games, such precision with respect to these formalisms rarely occurs. Instead, the term model regularly serves as the broadest, highest-level term of reference. For example, Gary Brewer and Martin Shubik performed a survey of wargame players, observers, developers, and funders, asking about the strengths and weaknesses of the games. This survey included detailed questions regarding the kind of formalism they worked with, used, developed, or funded. They summarized the survey responses as follows:

²⁵ Andrew Wilson, *The Bomb and the Computer: A Critical History of Wargames* (New York, NY: Delacorte Press, 1969).

We began by asking each respondent to identify his MSG (Model, Simulation, Game) as a model, a simulation, a man-machine or manual game, a mathematical analysis or study, or, if none of those categories fit, some “other” type...most were identified as models. The usage of that term and concept is vague, however; one respondent noted that to him a model was any regression equation, and the universe of models was therefore so large as to be virtually unbounded. In short, the word *model* has little referential utility. This terminological problem is underscored by the GAO’s [General Accounting Office’s] initial lack of success when it asked the military services to provide data on war games. The GAO found that practically no war games are played; rather models and simulations are built.²⁶

Understanding the differences between models, simulations, and games is important for developing and promoting the use of models in intelligence analysis. As Brewer and Shubik found, those who examined potential military conflicts through wargaming and simulations did not see their investigations as modeling, creating a disconnect between their believed non-use of models and actual employment of them. As a result, the importance of models in their many forms may go unrecognized, particularly in case of simulations and games.

Intelligence analysis has shown a similar pattern to the findings by Brewer and Shubik. Analysts have largely eschewed models, yet use simulations and games regularly in both formal and informal ways. As a result, many of the benefits and applications of models have gone unrecognized by the community, simply for the fact that practitioners often fail to recognize their relevance to intelligence problems and tradecraft. However, the practice of microscience and a model-centric tradecraft arise out of modeling practices already embedded in contemporary analytic tradecraft. Thus,

²⁶ Gary D. Brewer and Martin Shubik, *The War Game: A Critique of Military Problem Solving* (Cambridge, MA: Harvard University Press, 1979), pp. 161-163.

while models have been acknowledged in particular applications, most notably data mining, pattern recognition, and predictive analytics, the uses of simulations and games in the much richer and challenging context of strategic intelligence analysis and production has gone unrecognized.

Bounded Rationality

The uses of models in scientific inquiry and intelligence analysis are linked by the theory of bounded rationality and “thinking about thinking” that rests at the core of research on human cognition and problem solving. Bounded rationality was developed by Herbert Simon based on his research into decision making and management, particularly in government, business, and science.²⁷ Bounded rationality acknowledges that people possess innate limitations on the amount of information they can process and remember. As a result, how they choose to define or represent a problem, the representation’s resolution or detail, and the ways in which these representations are manipulated determine how solutions to problems are discovered and evaluated.

Because human information processing capabilities are limited and the costs of gathering new information and finding new representations can be expensive in terms of time and resources, people *satisfice* rather than optimize. They search for solutions that are good enough to suit their needs rather than find the best of all possible solutions simply because it is expensive or impossible to consider everything. Simon’s theory of bounded rationality provided the basis of Heuer’s *Psychology of Intelligence Analysis*,

²⁷ Herbert A. Simon, *Models of My Life* (Cambridge, MA: MIT Press, 1996); and Hunter Crowther-Heyck, *Herbert A. Simon: The Bounds of Reason in Modern America* (Baltimore, MD: Johns Hopkins University Press, 2005).

and the development of SATs that undergird contemporary analytic tradecraft, and also provided the foundation for many of the developments in modeling and simulation in the social sciences over the last two decades.²⁸

While bounded rationality has been acknowledged in intelligence analysis, arguments for its importance apply to all of science and problem solving more broadly. As Thomas Kuhn, John Ziman, and other philosophers have noted, science is not necessarily a rational act but is instead shaped by preexisting ideas and beliefs about what constitutes a question worthy of investigation, the shape or form of acceptable solutions, and the importance of individual psychology and group identity in determining the appropriateness of different research methods. For some, the rationality of science could be preserved by restraining the definition to science—to include data gathering and hypothesis testing, while leaving theorization, model building, and hypotheses generation as creative acts that rest outside of its purview. Indeed, this was the very argument made by Isaac Ben-Israel, who noted, “The creation of hypotheses is not part of the scientific method,” while arguing that intelligence analysis could be practiced in a fashion consistent with science.²⁹

Alternatively, Simon’s theory of bounded rationality argued that theorization and hypothesis generation, the creative, artistic aspects of scientific inquiry, were in fact

²⁸ Richards J. Heuer, Jr., *Psychology of Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 1999), pp. 1-6. Also see Charles A. Lave and James G. March, *An Introduction to Models in the Social Sciences* (Lanham, MD: University Press of America, 1993); Joshua M. Epstein, *Generative Social Science: Studies in Agent-Based Computational Modeling* (Princeton, NJ: Princeton University Press, 2006); and John H. Miller and Scott E. Page, *Complex Adaptive Systems: An Introduction to Computational Models of Social Life* (Princeton, NJ: Princeton University Press, 2007).

²⁹ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 700.

subject to the logical development, evaluation, and manipulation of problem solving heuristics and accessible to scientific observation and description. Psychology and cognitive science have examined the ways in which researchers' mental models operate by framing problems and employing heuristics for solving them that can produce theories, models, and hypotheses as part of an identifiable, evolutionary process. As a result, bounded rationality's emphasis on procedural rationality, i.e. the process of problem solving, allowed it to shed light on creative and intuitive mental processes that are invisible to the traditional concerns of substantive rationality. Simon described the difference between these two kinds of rationality within the overall context of bounded rationality as follows:

A theory of bounded rationality, then, will be as much concerned with procedural rationality, the quality of the processes of decision, as with substantive rationality, the quality of the outcome. To understand the former, one must have a theory of the psychology of the decision maker; to understand the latter, one needs have only a theory of the goal (the utility function) and the external environment.³⁰

By focusing on the ways problem solvers of all kinds—whether scientists, intelligence analysts, or policy makers—search to find solutions and evaluate their properties, the importance and roles of model development, design, and assessment become increasingly clear. Without models, the human brain, replete with its innate cognitive limitations, could not effectively parse the vast quantities of real observations and possible futures about the world in order to understand its properties or shape its behavior. Indeed, research on learning and problem solving has noted that bounded

³⁰ Herbert A. Simon, "Bounded Rationality in Social Science: Today and Tomorrow," *Mind & Society*, Vol. 1 (2000), p. 25.

rationality and the construction of models summarized the fundamental challenges and solutions to understanding how discoveries occurred.³¹ Therefore, models should be regarded as cognitive aids to thinking, serving simultaneously as inputs and outputs of problem solving.

The Limits of Mental Models

Models may be mental, sometimes referred to as informal or verbal models, or formal and encapsulated in graphical, physical, mathematical, or computational form. Mental models are the representations of systems that people hold in their heads. Mental models are particularly flexible and can express the widest possible range of ideas about the composition and behavior of systems. However, this flexibility comes at a price because of the inherent cognitive limitations of the brain.

As noted above, there are limitations to the complexity of the representations that the human brain can handle with precision before errors emerge and increasing simplification is needed, hence the basis of the theory of bounded rationality and the need for models. These errors occur when people reach conclusions that are not logically possible based on the assumptions embedded in their mental models, and result when models contain many moving parts or interactions over multiple periods of time, involve determining the value of quantitative variables, and more.

The evaluation of counterfactuals and scenarios, the fundamental problem of choosing between alternatives, is particularly dependent on the operation and exploitation

³¹ David Klahr, *Exploring Science: The Cognition and Development of Discovery Processes* (Cambridge, MA: MIT Press, 2002), p. xi.

of mental models because they involve deviations from empirically known and established cases. As a result, even when people wish to explore alternative outcomes, they may be overly conditioned and biased to replicate what is known and familiar.³² For example, John Mueller, an international relations scholar, performed an assessment of the effects of nuclear weapons on maintaining the political stability of Cold War by mentally replaying the major events of the latter half of the twentieth century in the absence of nuclear weapons in order to determine whether a conventional war between the US and the Soviet Union might have resulted. In this case, the analysis included the Cuban Missile Crisis, a logically impossible event that occurred precisely because of the existence of nuclear weapons, thus violating the conditions of his model.³³

Excessive simplification can lead to overly abstract mental models, making them difficult to interpret and communicate. Informal mental models are constructed and manipulated via the use of spoken and written language, often resulting in a trade-off between flexibility and precision. Vali Nasr's 2007 book, *The Shia Revival*, argued that Shia sectarian identities would play an increasingly important role in the politics of Muslim communities around the world.³⁴ Nasr's informal model can be interpreted many different ways because the precise goals of revival politics are abstract and open to

³² Several cognitive biases that are shaped by dramatic and recent personal experiences, as well as the intuitive association between familiarity and correct or positive outcomes, are discussed in Daniel Kahneman's *Thinking Fast and Slow* (New York, NY: Farrar, Straus and Giroux, 2011).

³³ See John Mueller, "The Essential Irrelevance of Nuclear Weapons," *International Security*, Vol. 13, No. 2 (Autumn 1988), p. 66-68; John Mueller, *Quiet Cataclysm: Reflections of the Recent Transformation of World Politics* (New York, NY: Harper Collins, 1995), pp. 71-73; and Lars-Erik Cederman, "Rerunning History: Counterfactual Simulation in World Politics," in Philip E. Tetlock and Aaron Belkin, eds., *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives* (Princeton, NJ: Princeton University Press, 1996), pp. 252-253.

³⁴ Vali Nasr, *The Shia Revival: How Politics within Islam Will Shape the Future* (New York, NY: W. W. Norton and Company, 2007).

interpretation in a variety of different social contexts—Shia groups around the world might engage in separatist movements, form national political parties, seek increased economic and social prominence in the country, or take other actions that may be locally consistent with the revival hypothesis but inconsistent with the actions of communities in other locations. In each case, the importance of sectarian issues and identities in politics may increase, consistent with Nasr’s overarching thesis, but vastly different futures for the communities in question may result and are not specified. As a result, Nasr’s readers may all interpret his argument differently, resulting in divergent measures, indicators, and assessments of Shia political successes and failures.

Regardless of how creative and carefully described, there is no guarantee that a mental model, once communicated, will produce the same behavior in one mind to the next, meaning that each user may imagine different behaviors or results. As John Miller and Scott Page noted:

An important feature of any theoretical tool is its trade-off between flexibility and precision. Flexibility occurs when the model can capture a wide class of behaviors; precision requires the elements of the model to be exactly defined.

One approach to achieving maximum flexibility is to use long verbal descriptions of the phenomena of interest. This tradition is well established in economics; for example, [Adam] Smith’s rather lengthy *Wealth of Nations* models a grand scheme of how the selfish behaviors of individuals can result in an outcome with surprisingly organized aggregate behavior. Such verbal descriptions, while flexible, often suffer from a lack of precision. There is an inherent ambiguity to such theorizing in terms of what is being expressed. More important, the implications of these types of descriptions are often difficult to verify—it is possible to make an apparently logical and coherent verbal argument that may in fact contain serious flaws.³⁵

³⁵ John H. Miller and Scott E. Page, *Complex Adaptive Systems: An Introduction to Computational Models of Social Life* (Princeton, NJ: Princeton University Press, 2007), pp. 78-79.

Thus, ambiguity and confusion often result where research and problem solving rely exclusively on mental models whose implications cannot always be determined intuitively. For this reason, many of the prominent debates within and between international relations theories have lasted for generations and show little signs of abatement or resolution. Thomas Cusack and Richard Stoll noted in 1990 that realism, a theory of international relations whose intellectual origins can be traced back to Thucydides, remained incapable of making consistent predictions and that its adherents remained divided regarding their expectations about the international system.³⁶ In the introduction to their book, *Exploring Realpolitik: Probing International Relations Theory with Computer Simulation*, Cusack and Stoll noted:

A great deal of conventional wisdom as well as scientific endeavor has been guided by realism. Still, it would be fair to say that neither is the theoretical structure implicit in the many writings within this tradition well understood, nor are the empirical expectations that follow from it fully confirmed. Our contention is that both of these problems can be reduced by focusing attention on the formalization of the ideas that exist within realism.³⁷

The empirical study of mental models notes that their users regularly employ simulation to solve problems. Specifically, they manipulate a version of the system in their minds, changing its state according to imagined process and interactions unfolding in time. However, for mental models to be dynamically manipulated, they must be quite simple in order to remain in the individual's short-term, working memory. As a result,

³⁶ Chris Brown and Kristen Ainley, *Understanding International Relations* (New York, NY: Palgrave MacMillan, 2009); and James E. Dougherty and Robert L. Pfaltzgraff, Jr., *Contenting Theories of International Relations: A Comprehensive Survey* (New York, NY: Pearson, 2009).

³⁷ Thomas R. Cusack and Richard J. Stoll, *Exploring Realpolitik: Probing International Relations Theory with Computer Simulation* (Boulder, CO: Lynne Rienner Publishers, 1990), p. 1.

the number of actors, their relative diversity and sophistication of behavior, the number of interactions that occur, and the total number of discrete time steps that can be considered are all limited. These points were made by Gary Kline, who studied how firemen thought about rescuing occupants in a car crash, and noted the extent to which many important features of the scenario, such as the type of car, circumstances of the crash, etc., were discarded as decision makers worked through the problem in their minds.

[T]he mental simulations were not very elaborate. Each seemed to rely on just a few factors—rarely more than three. It would be like designing a machine that had only three moving parts. Perhaps the limits of our working memory had to be taken into account. Also, there was another regularity: The mental simulations seemed to play out for around six different transition states, usually not much more than that. Perhaps this was a result of limited working memory. If you cannot keep track of endless transitions, it is better to make sure the mental simulation can be completed in approximately six steps.³⁸

The cognitive limitations of the human brain affect the complexity of mental models and how they can be manipulated. As problems become increasingly complex and may require sustained engagement in order to manipulate and assess the ability to understand the situation, conclusions reached through the analysis of mental models alone become suspect.

Placing mental models at the center of the problem solving also provides important insights into the meaning and value of expertise. A model-centric approach to scientific research and intelligence analysis allows for the distinction between experts and novices that is based on the sophistication (or elegance) of their mental models and the

³⁸ Gary Klein, *Sources of Power: How People Make Decisions* (Cambridge, MA: MIT Press, 1998), p. 52.

types of manipulations they can perform, rather than their predictive accuracy. This reinforces bounded rationality's emphasis on procedural vs. substantive rationality, where expertise is measured by the collection of heuristics and their manipulation. The linkage between prediction and procedures is then understood to be an evolutionary one, where experience and learning eliminate or downgrade mental models that consistently perform poorly.³⁹ In cases where the pruning of mental models is too aggressive, the result may induce analytic failures by experts, as new problems reveal that their repertoire of mental models are overly constrained by their paradigms, research programs, or mindsets that were too successful in eliminating rival models from consideration.

Another limitation of mental models is their ability to cope with games, where actors compete, cooperate, or coordinate. Most problems in strategic intelligence are games, where each actor has different information and goals, and must communicate, deceive, or keep secrets in order to achieve his or her goals. In these cases, the capability of the brain to partition itself and maintain private information and strategies of multiple actors is limited. For example, one cannot play a game of chess against oneself and keep black's and white's strategies, plans, and perceptions separate. As a result, analysts may need to employ formal models of games, e.g. game theory, or engage in simulations where individual analysts each play the role of different actors in the system.

...two or more centers of consciousness are dependent on each other in an essential way. Something has to be communicated; at least some spark of recognition must pass between players. There is generally a necessity for

³⁹ Herbert A. Simon, *The Sciences of the Artificial* (Cambridge, MA: MIT Press, 1996), pp. 85-110; Bent Flyvbjerg, *Making Social Science Matter: Why Social Inquiry Fails and How it Can Succeed Again* (New York, NY: Cambridge University Press, 2001), pp. 10-20; and Richard K. Betts, *Enemies of Intelligence: Knowledge & Power in American National Security* (New York, NY: Columbia University Press, 2007), pp. 53-65.

some social activity, however rudimentary or tacit it may be; and both players are dependent to some degree on the success of their social perception and interaction. Even two completely isolated individuals, who play with each other in absolute silence and without even knowing each other's identity, must tacitly reach some meeting of minds.

There is, consequently, no way that an analyst can reproduce the whole decision process either introspectively or by an axiomatic method. There is no way to build a model for the interaction of two or more decision units, with the behavior and expectation of those decision units being derived by purely formal deduction. An analyst can deduce the decisions of a single rational mind if he knows the criteria that govern the decisions; but he cannot infer by purely formal analysis what can pass between two centers of consciousness.⁴⁰

Modeling the interactions of more than one conscious actor requires formal techniques that can maintain the private information and secret strategies of each actor. Additional sources of error in mental modeling occur whenever systems contain feedback loops, experience exponential growth, contain time lags and delays between cause and effect, or possess other non-linear responses between inputs and outputs.⁴¹

Despite their limitations, however, mental models remain the ultimate arbiters of scientific and analytic judgment. Formal models provide valuable opportunities to rigorously examine ideas that reside in the human mind. However, the purpose of formalizing mental models is not to replace or offload judgment to external constructs, but to gain better insights into the systems under investigation in order to inform decision making at personal and organizational levels.⁴² Thus, formal modeling does not replace mental models or the expertise of their developers and users. Instead, formal modeling

⁴⁰ Thomas C. Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1980), p. 163.

⁴¹ Dietrich Dorner, *The Logic of Failure: Recognizing and Avoiding Error in Complex Situations* (Cambridge, MA: Perseus Books, 1996).

⁴² Hunter Crowther-Heyck, *Herbert A. Simon: The Bounds of Reason in Modern America* (Baltimore, MD: Johns Hopkins University Press, 2005), p. 102.

provides users with new information based on representations and manipulations that cannot be reliably performed by the unassisted brain—information that must be evaluated, assimilated, and ultimately filtered back through mental models.

Types of Formal Models

Formal models are the externalized representation and construction of mental models and the systems they represent. In order to cope with the bounded rationality, researchers have increasingly employed external aids to compensate for their cognitive limitations. Such aids might be as simple as writing down a phone number or email address because it cannot be easily remembered, to algorithmic trading applications that can recognize and act on patterns in financial data faster and more precisely than human beings. In each case, humans have formalized and externalized their mental models in order to accomplish what they cannot by relying on their innate cognitive faculties alone. This section provides a summary of the major categories of formal models and their relative strengths and weaknesses. The different types of formal models discussed below are physical, mathematical, computational, or graphical.

Physical Models

Physical models, often called scale models, are tangible representations of systems used to study their properties through manual inspection and manipulation. These models usually resemble the physical aspects of the system they are representing, such as a model train, airplane, or submarine. Physical models may also represent the properties of abstract systems through the exploitation of physical mechanisms, such as

economic models employed in the early and mid-twentieth century that studied the flow of money, interest rates, credit, debt, etc. through the use of water, pumps, pipes, and containers.⁴³ Physical models may also represent real-world systems that are physical, but ultimately inaccessible to researchers, including models of the solar system, atoms, refracting light, and more.⁴⁴

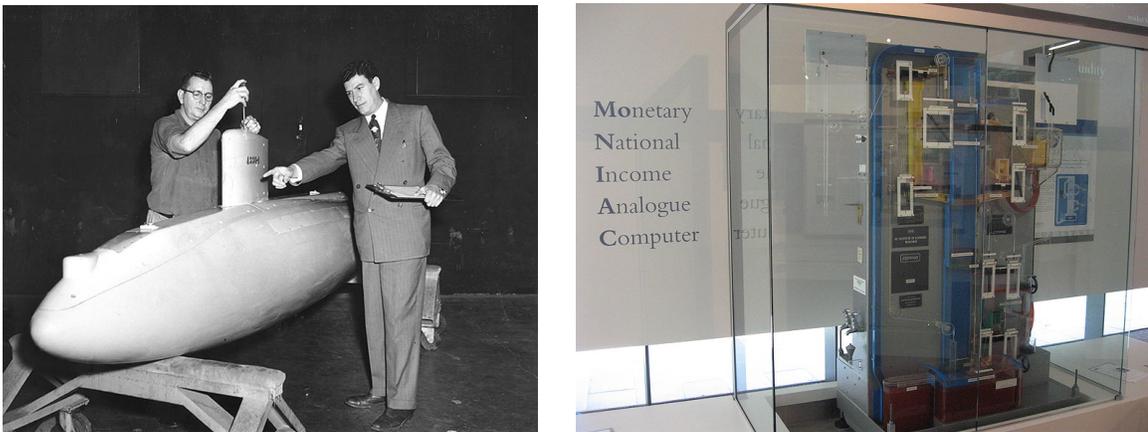


Figure 2-1: Physical Models. The left image shows a model of the USS Albacore being prepared for tests at the David Taylor Model Basin in 1956. On the right is a depiction of the Monetary National Income Analogue Computer (MONIAC) analytic computer, created in 1949, that represented the national economy using colored water and pumps to simulate flows of money to different economic sectors. Images from Naval Historical Records Center, “USS Albacore (AGSS-569), 1953-1984,” <http://www.history.navy.mil/photos/sh-usn/usnsh-a/ss569.htm> (accessed on August 5, 2011) and Wikipedia, “MONIAC Computer,” <http://en.wikipedia.org/wiki/File:MONIAC.jpg> (accessed August 5, 2011).

Historically, physical or scale models used to be referred to as mechanical analogies, which offers important insights into their use in scientific research. While

⁴³ Wikipedia, “MONIAC Computer,” http://en.wikipedia.org/wiki/MONIAC_Computer (accessed August 5, 2011).

⁴⁴ Daniela Bailer-Jones, *Scientific Models in Philosophy of Science* (Pittsburgh, PA: University of Pittsburgh Press, 2009), pp. 21-44.

their employment for the purposes of illustration or demonstration is relatively straightforward, their use in actual research is more complex because model building at different scales and materials necessarily involved using one system to represent another system, e.g. the use of billiard balls to represent the interactions of molecules in a gas. This gave rise to the notions of negative, positive, and neutral analogies as a way of identifying the relationships between the constructed model and the target system. As Mary Hesse noted, negative analogies are those properties of models that do not accurately represent, or distort, the properties of the target system. Positive analogies are those properties that are shared between the model and the target system. Neutral analogies are those properties of the model that are not yet known to be similar or different from the target system.⁴⁵

The notion of positive, negative, and neutral analogy with respect to the status of models is an important one when considering how to make inferences from the system that is accessible about one that is not. For example, during the rehearsals of the Abbottabad raid, a physical model of the compound was constructed. However, this model was partial—the positive analogy referred to the size and shape of the compound, as well as many of its geographic features such as weather and altitude.⁴⁶ Alternatively, the neutral analogy included properties such as the internal layout of the compound's buildings, gates, and security devices. This was evident during the raid, where the SEAL team had planned how to approach and enter the compound, but was forced to improvise

⁴⁵ Daniela Bailer-Jones, *Scientific Models in Philosophy of Science* (Pittsburgh, PA: University of Pittsburgh Press, 2009), pp. 54-55.

⁴⁶ Peter L. Bergen, *Manhunt: The Ten-Year Search for Bin Laden—From 9/11 to Abbottabad* (New York, NY: Crown, 2012), p. 184.

tactically once inside because they did not know about the building's interior layout or security features.

The SEAL team had no idea what the layout of the floors inside bin Laden's house might be. As they moved deeper inside, they passed a kitchen and two large storage rooms. Near the back of the house, which had a bunker-like feel, was a stairwell. Blocking their way to the upper two floors was a massive, locked metal gate. The SEALs blasted their way through this gate with the breaching materials they were carrying.

Leiter says that he was concerned that the house might be booby-trapped with bombs, a technique al-Qaeda had perfected in Iraq: "I kept waiting for some big explosion from the house that just made everything sink." Brennan was also anxious: "Might there be a quick reaction force that bin Laden may have had, security that we didn't know about?"⁴⁷

Alternatively, the negative analogy proved to be the most important factor in the conduct of the SEALs' raid. Small differences in building materials, particularly the use of chain-link fence in the model that represented where the real-world compound had a concrete wall, affected the aerodynamics of the approaching helicopters, causing one to crash.

The Black Hawks approached Abbottabad from the northwest. Once the helicopters reached their destination, the carefully planned operation began to unravel. As the first chopper tried to land in the largest courtyard in the compound, it suddenly lost altitude. The combination of the additional weight of the stealth technology and the higher-than-expected temperatures in Abbottabad had degraded its performance, causing an aerodynamic phenomenon known as "settling with power," meaning an unexpectedly fast drop. When the SEALs had practiced the maneuver on a replica of the compound in the States, the compound's outer walls had been represented by chain-link fencing, whereas the actual walls were made of concrete. The thick walls likely gave more energy to the Black Hawk's rotor wash and contributed to the chopper's instability. Because of that instability, the tail of the craft clipped one of the compound walls, breaking off the critical tail rotor....

⁴⁷ Peter L. Bergen, *Manhunt: The Ten-Year Search for Bin Laden—From 9/11 to Abbottabad* (New York, NY: Crown, 2012), p. 223.

...Over the course of half an hour, in a small classroom on the base, the president was briefed by the men who had carried out the operation....

The SEAL team ground commander used a model of the compound and a red laser pointer to explain what had gone right and what had gone wrong on the mission from start to finish. The biggest problem was not correctly matching the wall that surrounded the compound in the life-size mock-ups they had used for rehearsals. The solid walls of the actual compound had caused turbulent aerodynamics for the first Black Hawk when it hovered to drop the SEALs into the courtyard, and had necessitated the chopper's "hard landing."⁴⁸

Physical or scale models are not always manually constructed. In some cases, models may be living systems—organisms or ecosystems selected for study and experimentation in order to reveal new insights about other systems that are less accessible. This is often the case in biology and medicine, where one species is examined in order to make inferences about another. In these situations, one system is treated as a surrogate for another.

Unlike the idealized representational models characteristically featured in the history of the exact sciences, in which the model (e.g. the Bohr atom) has been supposed to mirror a natural system (hydrogen) by embodying the mathematical laws and structure from which the behavior of the system can be deduced, model systems maintain their own autonomy and specificity. That is, model systems do not directly represent humans as models *of* them. Rather, they serve as exemplars or analogues that are probed and manipulated in the search for generic (and genetic) relationships. They serve as models *for* human attributes. The use of standardized organisms in biomedicine is part of a broader model-systems approach in the life sciences that includes the investigation of a far wider range of entities, from specific proteins (e.g. hemoglobin) to particular lakes (e.g., Linsley Pond in Connecticut), and whose utility in producing general knowledge relies on the routine use of analogies to other examples and entities.⁴⁹

⁴⁸ Peter L. Bergen, *Manhunt: The Ten-Year Search for Bin Laden—From 9/11 to Abbottabad* (New York, NY: Crown, 2012), pp. 218-219, and 247.

⁴⁹ Angela N. H. Creager, Elizabeth Lunbeck, and M. Norton Wise, "Introduction," in Angela N. H. Creager, Elizabeth Lunbeck, and M. Norton Wise, eds., *Science Without Laws: Model Systems, Cases, Exemplary Narratives* (Durham, NC: Duke University Press, 2007), p. 2.

Mathematical Models

Mathematical models are representations of systems as equations. Mathematical models have been used in every scientific discipline and have been applied to a wide variety of problems ranging from armed combat to economic trends to environmental sustainability and pandemics. In general, mathematical models exist in two forms. They may be statistical, in which case models are generated from data via induction, or deductive representations of assumptions or conjectures about how a system works.

Statistical models describe patterns and relationships in data, providing information about the distribution or structure of variables within a dataset, usually resulting in measures of correlation or assessments of the probability of future events based on the historical occurrences of past events. Inductive models are often used to test hypotheses in order to determine if changes in an independent variable have a statistically significant effect on a dependent variable. Statistical models vary greatly with respect to the sophistication of the mathematics they employ and the ways in which they search for patterns in data, but are limited by the availability and reliability of appropriate empirical data. Thus, statistical approaches are most useful in identifying patterns and structures in the data analysts possess, but face difficulties when used to explore future scenarios beyond projecting existing patterns forward in time.

Deductive models express their developer's assumptions about a system in the language of equations in order to study their logical consequences. These models are not statements about the empirical world, as is in the case with statistical models, but are themselves hypotheses or conjectures that reveal the implications of chosen assumptions

through numerical or symbolic manipulation. Because deductive models are based on imagination rather than data, they are only bounded by the limits of mathematical expression and manipulation: If an assumption about a system can be expressed as an equation it can be modeled mathematically, though not always solved (in such cases the properties can be explored through numerical simulation, a special application of computation to mathematics). This allows for deductive models to look beyond the limitations of available data, and engage in speculative analysis as a form of mathematical experimentation—a series of “what if” investigations explored through equations.

Proponents of mathematical models have argued that the language of mathematics is more precise than those of informal, verbal models.

A pragmatic reason is that it simply allows us to communicate with each other in an orderly and systematic way; that is, ideas expressed mathematically can be more carefully defined and more directly communicated than with narrative language, which is more susceptible to vagueness and misinterpretation. The causes of these effects include multiple interpretations of words and phrases, near-synonyms, cultural effects, and even poor writing.⁵⁰

The benefits of precision allow for model users to refine their research and focus their attention on aspects of problems that might otherwise be overlooked or remain unspecified. Essential to this perspective is the notion that models are not predictors of future outcomes, but rather investigative tools that provide insights to their users.

Because models are simplified representations of systems, they never tell the entire story of how it works, and what they get wrong may be just as important—and often more

⁵⁰ Jeff Gill, *Essential Mathematics for Political and Social Research* (New York, NY: Cambridge University Press, 2006), pp. 1-2.

important—than what they get right when compared with target systems in the real world.

An example of this can be seen in a simple mathematical model of voter turnout.⁵¹ This model consists of three independent variables that calculate the utility of voter participation in an election. These variables are:

- the effect of the individual's vote on the electoral outcome;
- the voter's perceived difference between the candidates and the policies they would implement on matters of importance to the voter; and
- the costs of participating in the election to the voter, taking time, energy, money, etc. from other activities the individual could perform.

The logic of the model can be expressed as follows:

The utility of voting is determined by the probability that an individual's vote will determine the outcome of the election and the perceived difference between the candidates minus the cost of participation.

A more compact representation of the logic is the equation:

$$R = PB - C$$

- R is the utility of voting to the individual voter
- P is the probability of the individual's vote affecting the outcome of the election
- B is the perceived difference between the candidates being chosen
- C is the cost of voting with respect to time, effort, money, etc.

Examination of the model's behavior leads to several conclusions. Voters will not participate if the costs of voting (C) exceed the perceived value of their vote when deciding between the different candidates (PB). If voters perceive the difference between the candidates to be small, the utility of voting is diminished when compared to cases

⁵¹ Jeff Gill, *Essential Mathematics for Political and Social Research* (New York, NY: Cambridge University Press, 2006), pp. 5-6.

where the perception is a large difference in their platforms and policy goals. Voters may not participate in an election if other potential activities can produce a higher utility (R). As the expected turnout increases, the perceived value of the individual's vote declines, making participation less likely (P). These insights lead to a "voter's paradox" where the value of individual votes declines as the number of voters increases, yet large turnout for elections nevertheless occurs.

The "voter's paradox" is an example of a partial solution derived from a formal model that shifts the character of the questions researchers ask. For example, the model does not include other socially relevant variables, such as identity or social norms that transform electoral participation into a civic responsibility as opposed to treating it as an individual utilitarian calculation. Ironically, their absence from the model, and peculiar results such as the underestimation of voter turnout, reveals the importance of those variables that are missing from the model. Thus, the model's utility is not determined by its predictive accuracy, but by its ability to expose the relative value of different variables and prospective relationships between them.

Mathematical models play a privileged role in the history of science, and by extension, world history.⁵² The ability to express ideas as mathematical constructs enabled the communication, exploration, and testing of theories against experimental and observational data, i.e. the core components of the scientific method. The importance of mathematical models has even been cited as a central force behind Europe's distinct economic and military expansion during the Renaissance, often referred to as the "Rise of

⁵² Karl Popper, *The Poverty of Historicism* (New York, NY: Routledge, 2004).

the West,” “Military Revolution,” or “Great Divergence.”⁵³ These developments rested upon an intellectual foundation and devotion to measurement and the search for reliable, exploitable knowledge out of which new sciences, engineering, and business practices emerged.

The distinctiveness of the European mindset with respect to mathematical modeling was noted by Arnold Pacey who argued that Europeans imbued technologies with symbolic power, seeing clocks, the printing press, and other mechanical devices as models of an orderly, predictable, and controllable universe. Thus, cultural differences about how to interpret technologies, and their role as models of larger, fundamental principles stimulated the search for new knowledge—creating a self-reinforcing feedback loop of discovery that distinguished Europe from the dominant global powers of the sixteenth century in the Islamic world and China.

One symptom of this difference in interests is attitudes to clocks. In the West they were seen as of great significance in showing complex motions of the planets reduced to a model with simple parts. However, there is a difference between *models* as straightforward representations of material objects, and *symbols* as signs of profounder realities. Seen as a model, the clock encouraged the construction of other mechanical models, some of them merely clockwork toys, made to entertain. Seen as a symbol, the clock represented the growing belief that a form of *mechanical order* permeated the whole universe. This encouraged people to conceptualize natural processes in terms of mathematical relationships known to be applicable to machines such as clocks.⁵⁴

⁵³ William H. McNeill, *The Rise of the West: A History of Human Community* (Chicago, IL: University of Chicago Press, 1991); Geoffrey Parker, *The Military Revolution: Military Innovation and the Rise of the West, 1500-1800* (New York, NY: Cambridge University Press, 1996); and Kenneth Pomeranz, *The Great Divergence: China, Europe, and the Making of the Modern World Economy* (Princeton, NJ: Princeton University Press, 2000).

⁵⁴ Arnold Pacey, *Technology in World Civilization: A Thousand-Year History* (Cambridge, MA: MIT Press, 1991), pp. 96.

Pacey's assessment regarding the cultural differences in how European and other cultures viewed technology was further elaborated on by Alfred Crosby, who reflected on the rise of European states after the Middle Ages and concluded:

I was nagged by the impression that Europeans were incomparably successful at sending ships across oceans to predetermined destinations and at arriving at those destinations with superior weaponry—with, for instance, cannons superior to those of the Ottomans and the Chinese; that they were more efficient at operating joint-stock companies and empires of unprecedented extension and degree of activity than anyone else—that they were in general far more effective than they have been, at least as judged by their own and other's precedents. Europeans were not as magnificent as they believed, but they were able to organize large collections of people and capital to exploit physical reality for useful knowledge and for power more efficiently than another people of the time. Why?

The textbook answer is, put simply, science and technology, and that was certainly true for generations and still is in large parts of the world. But if we gaze back through and beyond the nineteenth century to the beginnings of European imperialism, we see little science and technology as such. Westerners' advantage, I believe, lay at first not in their science and technology, but in the utilization of habits of thought that would *in time* enable them to advance swiftly in science and technology and, in the meantime, gave them decisively important administrative, commercial, navigational, industrial, and military skills. The initial European advantage lay in what French historians have called *mentalite*.

During the late Middle Ages and Renaissance a new model of reality emerged in Europe. A quantitative model was just beginning to displace the ancient qualitative models. Copernicus and Galileo, the artisans who taught themselves to make one good cannon after another, the cartographers who mapped the coasts of newly contacted lands, the bureaucrats and entrepreneurs who managed the new empires and East and West India companies, the bankers who marshaled and controlled the streams of new wealth—these people were thinking of reality in quantitative terms with greater consistency than any other members of their species.⁵⁵

⁵⁵ Alfred W. Crosby, *The Measure of Reality: Quantification and Western Society, 1250-1600* (New York, NY: Cambridge University Press, 1997), pp. x-xi.

Historically, mathematical equations used to be regarded as something distinct from models, portraying the true state of systems, whereas physical models were considered analogies. Physical models were employed whenever the underlying, true mathematical structure of a system was unknown or too complex to be manipulated, thus elevating mathematical models to a more sophisticated or mature form of knowledge about a phenomenon. However, during the twentieth century the term model expanded in its meaning as equations were increasingly acknowledged to be a type of uncertain representation of problems akin to physical models once it was demonstrated that the same phenomena could be faithfully represented by different equations with divergent implications. Such a change brought mathematical equations into the same tent as physical models: Models were no longer regarded as objective statements of physical laws but as theoretical abstractions with properties of their own that impinge upon the systems they represent.

Throughout the course of the twentieth century, consideration of scientific models has nonetheless been resumed, certainly by scientists and also by philosophers of science. This became possible partly because the concept of model changed. Models were no longer seen as having to be mechanical. Bohr's model of the atom, for instance, is both highly abstract and mathematical. Although some models have simply ceased to be mechanical, for others what it means to be mechanical has changed. For example, Rom Harre...has specifically explored the link between models and mechanisms in an article called "Metaphor, Model, and Mechanism." The same text reappeared in *Theories and Things*...entitled "Models to Mechanisms." Harre begins with what he calls principle P₁: "If you don't know why certain things happen then invent a mechanism (in accordance with the view you take of how the world works)—but it is better still if you find out how nature really works...."

What scientists do is not necessarily discover mechanisms, but check hypotheses about mechanisms.... Harre calls these hypothetical mechanisms and equates them with models. It is a characteristic feature of science to accept or reject a hypothetical mechanism, or model, as the real

mechanism.... Hypothetical mechanisms are necessary “because we have reached the limit of discernible mechanisms....” This is precisely a reference to the way science has changed and to the way the concept of a mechanism may be changed with it.⁵⁶

Computational Models

Computational models were originally viewed as extensions of mathematical models, allowing for the use of datasets, and solving equations faster than were more complicated than could be performed manually.⁵⁷ However, computational models evolved into distinct formalisms that emphasized the representation of systems as algorithms rather than equations. Whereas mathematical models are solved, computational models represent systems as processes that are executed as a set of instructions or rules that dictate how variables interact or actors behave.⁵⁸ This means that dynamics that are difficult to represent as equations, such as the training regimen of foreign military forces, the construction of Improvised Explosive Devices (IEDs) by Iraqi insurgents, or processes of recruitment into violent, radical extremist groups may be represented as a series of rules that are closer approximations to the behavioral descriptions of social, technical, or organizational processes that analysts employ. Because of their differences, computational models offer many new opportunities for modeling social systems and intelligence challenges that have resisted mathematical examination.

⁵⁶ Daniela M. Bailer-Jones, *Scientific Models in Philosophy of Science* (Pittsburgh, PA: Pittsburgh University Press, 2009), p. 36.

⁵⁷ Charles S. Taber and Richard J. Timpone, *Computational Modeling* (Thousand Oaks, CA: SAGE Publications, 1996), pp. 1-2.

⁵⁸ H. Van Dyke Parunak, Robert Savit, and Rick L. Riolo, “Agent-Based Modeling Vs. Equation-Based Modeling: A Case Study and User’s Guide, *Proceedings of Workshop on Modeling Agent Based Systems*, Paris, 1998.

Computational models are not limited to representing systems as equations and reducing variables to numerical terms. Instead, they may employ sequences of conditional steps and decisions, allowing for the examination of decision-making processes based on how actors gather and evaluate information in order to make choices based on a series of “if-then” conditions. As a result, computational models distinguish themselves by their ability to simulate processes, which provide greater flexibility in representing real-world decision making. This is particularly relevant within the social sciences, because it enables simulation models to relax many of the assumptions embedded in mathematical models of human and animal behavior with respect to rationality, geography, and diversity. By emphasizing logical processing, computational models enable the representation of systems to transition from substantive rationality and its emphasis on outcomes, to procedural rationality and its emphasis on individual and group decision-making processes.

The ability to simulate bounded rational agents that employ decision-making processes of arbitrary complexity, from “zero-intelligence” to extremely rich cognitive processes, has created new research opportunities to study complex systems.⁵⁹ While models of all kinds aid their users in addressing their own innate cognitive limits, computational models enable the actors in modeled systems to be represented as bounded rationality as well. As a result, computational models allow for the relaxation of many assumptions embedded in mathematical models, particularly with respect to rationality,

⁵⁹ Dhananjay K. Gode and Shyam Sunder, “Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality,” *The Journal of Political Economy*, Vol. 101, No. 1 (February 1993), pp. 119-137; and John E. Laird, *the SOAR Cognitive Architecture* (Cambridge, MA: MIT Press, 2012).

providing new opportunities to employ formal models in the study of complex systems and problems.⁶⁰ Indeed, the development of cognitive science that seeks to understand the behavior of agents based on their internal mental processes as a type of computation is regarded as one of the most significant developments in science over the last few decades.

Possibly the most significant event of the last 30 years is the rise of cognitive science, which studies various aspects of human cognition such as perception, memory, learning, and reasoning, and has transformed traditional psychology. Much of the impetus for cognitive science comes from the idea that the human mind is in some respects similar to a computer, and thus that human mental processes can be understood by comparing them to the operations computers carry out.⁶¹

ABM constitutes one of the many different approaches to computational modeling and is most commonly employed for simulating complex adaptive systems where the actions and interactions of multiple actors compel adaptation, learning, and evolution. ABM is distinctive when compared with other computational modeling formalisms because it deliberately incorporates social theory into the specification and behavior of the agents. In other cases, computational models are actually mechanisms that allow for the manipulation of mathematical, physical, or even other computational models in order to search for solutions to problems such as parameter estimation or optimization via

⁶⁰ Herbert A. Simon, "Bounded Rationality in Social Science: Today and Tomorrow," *Mind & Society*, Vol. 1, No. 1 (2000), pp. 25-39.

⁶¹ Samir Okasha, *Philosophy of Science: A Very Short Introduction* (New York, NY: Oxford University Press, 2002), pp. 10-11.

techniques that could not be manually performed, e.g. simulated annealing or genetic algorithms.⁶²

Visual Models

A fourth type of formal model is visual, sometimes called graphical or organizational. These models arrange information for the purposes of conceptualization, manipulation, and communication. Visual models may be as simple as the use of tables, matrices, Venn diagrams, flow charts, network diagrams, and other representations of systems. Many of the SATs that have been developed for helping intelligence analysts ensure that they have articulated their assumptions and explored alternative hypotheses focus on employing visual models, such as the development of complexity maps, quadrant crunching, and Alternative Competing Hypotheses matrices.⁶³

Within the intelligence community, Sherman Kent's pyramid of intelligence types is one important visual model, while the intelligence cycle is another commonly used model, shown in Figure 2-2. These visual depictions of the intelligence cycle show important differences in the role of consumers in each version, but also reveal that interpreting visual models can require a specialized set of skills. For example, the differences between the two versions of the intelligence cycle are subtle, and only apparent to viewers who are aware of the persistent tensions that exist between producers and consumers.

⁶² Melanie Mitchell, *An Introduction to Genetic Algorithms* (Cambridge, MA: MIT Press, 1996); and Daniela M. Bailer-Jones, *Scientific Models in Philosophy of Science* (Pittsburgh, PA: Pittsburgh University Press, 2009), pp. 63-64.

⁶³ Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011).

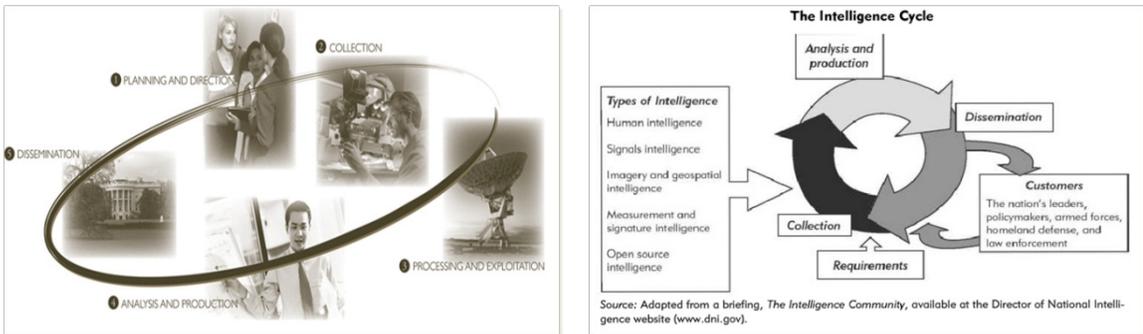


Figure 2-2: Two Versions of the Intelligence Cycle. Two versions of the intelligence cycle are depicted above. On the left is the standard version that shows the high-level steps associated with intelligence tasking, collection, analysis, and dissemination. On the right is a slightly modified form that has one significant difference: Consumers sit outside of the process, indicating that the intelligence community may execute its tasks without being dependent on consumers for setting requirements or accepting finished analytic products once disseminated. Images from the Central Intelligence Agency, *The Intelligence Cycle*, <https://www.cia.gov/library/publications/additional-publications/the-work-of-a-nation/work-of-the-cia.html> (accessed on June 30, 2012); and James B. Bruce and Roger Z. George, “Intelligence Analysis—The Emergence of a Discipline” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence* (Washington, DC: Georgetown University Press, 2008), p. 12.

In one version of the intelligence cycle, policy makers are a part of the intelligence process, and it cannot be faithfully executed without their active participation in the planning and direction and in the receipt of disseminated finished intelligence products. In the alternate version, however, these same consumers sit outside of the cycle that runs perpetually within the intelligence community. Instead, policy makers inject their priorities into the intelligence community and extract analytic products, but the intelligence community itself functions autonomously and continuously in the absence of consumer input or feedback.

Often visual models encompass aspects of other modeling types. For example, the visualization of social networks is both a graphical product of Social Network Analysis (SNA), and employs several mathematical or computational analogies as part of its layout managers, such as the use of gravity or magnetic metaphors of attraction and repulsion between nodes in order to identify and layout clusters or components.

In summary, whether a formal model is physical, mathematical, computational, or visual, they are all externalized representations of their developers' mental models. Each model type exposes the inner workings of their creators' mental models. These mental models may appear as assumptions within the model with respect to system boundaries and components, level of analysis, or behavioral or physical processes and descriptions of interactions. They may also be more subtle, such as the selection and coding of data or the acceptance of mathematical assumptions embedded in statistical procedures. The skilled use of formal models requires researchers to tease out what results are directly attributable to the model's assumptions, and input data from those that are a result of particular methods, representational choices, or artifacts.

The Uses of Models

Models can be employed to meet a variety of needs on behalf of their users. Models are most commonly associated with prediction but have many other uses that can vary based on the contexts in which they are employed. Some uses primarily benefit the individual or small groups of researchers focused on intellectual discovery, while others are group oriented and embody processes, procedures, or even politics. Nine different uses of models are:

- Prediction
- Explanation
- Exploration/Hypothesizing
- Partial Analysis
- Organization
- Collaboration
- Illustration
- Management
- Training⁶⁴

In each of these cases, applications of models to the problems facing intelligence analysts and the community are discussed. Importantly, this expands the use of models for prediction in order to consider many other individual, group, and organizational considerations. In each case, ABMs may be used to address existing gaps or shortcomings in intelligence analytic and managerial processes.

Prediction

Prediction uses models to estimate the state of a system or value of a variable based on the initial conditions specified in a model. When a system is closed and stationary, where no new actors or information enter into the system and processes linking variables within the model do not change over time, then models can be very effective in projecting future outcomes. However, such cases are relatively rare, because real-world systems are open—new actors and information continuously enter the system and constantly affect the composition and behavior of system’s units.⁶⁵ As Andrew Sayer noted, the differences between open and closed systems are particularly important

⁶⁴ This list elaborates model uses discussed in James S. Hodges, “Six (Or So) Things You Can Do with a Bad Model,” *Operations Research*, Vol. 39, No. 3 (May-June, 1991), pp. 355-365.

⁶⁵ Andrew Sayer, *Method in Social Science* (New York, NY: Routledge, 1992), pp. 130-138.

for the social sciences, and establishing control, even if only fleeting, over real-world systems.

The social sciences deal with open systems but lack the advantage of their equivalences in natural science of having relevant closed system sciences on which to draw. One of the main reasons for the openness of social systems is the fact that we can interpret the same marital conditions and statements in different ways and hence learn new ways of responding, so that effectively we become different kinds of people. Human actions characteristically modify the configuration of systems, thereby violating the extrinsic conditions for closure, while our capacity for learning and self-change violates the intrinsic condition. Paradoxically, it is because most systems are open, and many relations contingent, that we can intervene in the world and create closed (non-human) systems. At the most, social systems can only be quasi-closed, producing regularities that are only approximate and spatially and temporally restricted. A considerable part of human labour and communication is devoted to the creation of closed or quasi-open systems, with the aim of taking advantage of and controlling mechanisms of value to us, be it photosynthesis in edible plants or the synchronization of labour in a factory. Many forms of social organization tend to produce approximate regularities in patterns of events by enforcing rules or by subordinating workers to machines, which routinize and control the spacing and timing of particular kinds of action. The conditions for closure are therefore of practical as well as academic importance.⁶⁶

Predictions usually result from the extrapolation of patterns in data identified by statistical models. For example, statistical analysis may be used to estimate the probability of future wars between rival states or changes in the price of oil. In each case, knowledge of prior trends and the distributional properties of known populations can be estimated, allowing for the prediction of future conditions and imputation of missing data (which is itself a prediction of what the value of the elements outside of a sampled population might look like). Statistical models are generally used to test hypotheses because they measure or estimate how a change in one variable affects another, but such

⁶⁶ Andrew Sayer, *Method in Social Science* (New York, NY: Routledge, 1992), pp. 123-124.

predictions are non-explanatory because they rely on measures of correlation rather than claims of causation.⁶⁷

Predictions may also be made on the basis of other types of mathematical, computational, or other formal models that emphasize causal processes and social interactions, e.g. game theory. In these cases, predictions may contain explanations because they provide a causal basis or mechanism that explains why a future state is expected to emerge. However, because social systems and the challenges intelligence analysts face involve strategic interaction between actors, fundamental epistemological limits exist regarding whether prediction is a viable use of models on the problems they study.

The first 50 years of the last century, perhaps even the previous 200 years, was dominated by the notion that science would yield answers of the simplest kind to a wide range of applicable problems but this certainty has gradually dissolved. The reasons for this are diverse. At one level, this may be no more than one of those unfathomable psychological shifts in our awareness of the limits to our knowledge which occur periodically; at another level, it may be due to an increasing body of experiential knowledge of using science in the quest for exact answers to important problems and the growing realization that such certainty is illusory. The recent history of social forecasting in this regard has been salutary; both macro and micro events, from predictions of the stock market and the general performance of the economy to more local issues such as demographic change and traffic movements in cities seem beyond our understanding as well as control in that extraneous events now seem to dominate their behavior. Although this may always have been the case, the models that were fashioned a generation or more ago now seem wholly inadequate.

None of this has daunted our curiosity in using science to explain and predict but it has changed it. 50 years ago, the quest to build useful theories and models was dominated by the view that we could simplify and distill the essence of things so that we might capture enough of the

⁶⁷ Kenneth N. Waltz, *Theory of International Politics* (Reading, MA: McGraw Hill, 1979), pp. 2-3; and Alexander Rosenberg, *Philosophy of Social Science* (Boulder, CO: Westview Press, 2007), pp. 8-10.

social reality sufficient for comprehension and decision. Despite recognition that the world was complex, it appeared simple enough to produce some absolutism to the theory and models that might be employed in applications. With this growth in uncertainty and the increasing perception that the systems that we deal with are intrinsically complex, simplicity no longer seems the watchword in the development of techniques and models. Prediction is couched in qualification, and our science has become less orientated to prediction but more an aid to understanding, to structure debate. This is seen nowhere more clearly than in the shift to constructing “what if” scenarios which now dominate all model-building.⁶⁸

As Michael Batty and Paul Torrens noted, the increasing complexity, or acknowledgement of its presence, has fundamentally shifted the use of models in the study of social systems away from prediction toward the role of explanation.

Explanation

Explanation is another use of models. Whereas prediction focuses on estimating the future state of a system, explanation seeks to provide model users with an understanding of how the system works. Sometimes explanation can facilitate a prediction, as is often the case with simulations and games that expose the dynamic forces acting within a system. In most cases, however, explanations are developed based on the study of history and the tracing of causal processes backward in time, starting from the end and working through preceding states and decisions, or in the theorization of mechanisms that produce particular outcomes.

The nature of explanation raises two important points that are rarely acknowledged and often not fully appreciated. First, because explanations often reason backward, the outcome of a specified process must be known in advance. Thus, they

⁶⁸ Michael Batty and Paul M. Torrens, “Modeling Complexity: The Limits to Prediction,” *University College of London Working Paper Series*, No. 36 (October 2001), p. 3.

cannot explain future events that have not occurred. Additionally, when looking backward, they are highly vulnerable to hindsight bias, where particular events appear more predictable in retrospect than they were contemporaneously, which not only affect the selection of models used to explain events, but also the expertise and wisdom of model users themselves. Khaneman noted the ways in which hindsight can affect how social situations are assessed, particularly those that intelligence analysts confront.

Hindsight bias has pernicious effects on the evaluations of decision makers. It leads observers to assess the quality of a decision not by whether the process was sound but by whether its outcome was good or bad. Consider a low-risk surgical intervention in which an unpredictable accident occurred that caused the patient's death. The jury will be prone to believe, after the fact, that the operation was actually risky and that the doctor who ordered it should have known better. This outcome bias makes it almost impossible to evaluate a decision properly—in terms of the beliefs that were reasonable when the decision was made.

Hindsight is especially unkind to decision makers who act as agents for others—physicians, financial advisers, third-base coaches, CEOs, social workers, diplomats, politicians. We are prone to blame decision makers for good decisions that worked out badly and to give them too little credit for successful moves that appear obvious only after the fact. There is a clear outcome bias. When the outcomes are bad, the clients often blame their agents for not seeing the handwriting on the wall—forgetting that it was written in invisible ink that became legible only afterward. Actions that seemed prudent in foresight can look irresponsibly negligent in hindsight....

The worse the consequence, the greater the hindsight bias. In the case of a catastrophe, such as 9/11, we are especially ready to believe that the officials who failed to anticipate it were negligent or blind. On July 10, 2001, the Central Intelligence Agency obtained information that al-Qaeda might be planning a major attack against the United States. George Tenet, director of the CIA, brought the information not to President George W. Bush but to National Security Adviser Condoleezza Rice. When the facts later emerged, Ben Bradlee, the legendary executive editor of *The Washington Post*, declared, "It seems to me elementary that if you've got the story that's going to dominate history you might as well go right to the president." But on July 10, no one knew—or could have

known—that this tidbit of intelligence would turn out to dominate history.⁶⁹

A second problem of explanation is that because the brain naturally infers causation, it invents explanations for events that may be the result of purely stochastic processes. If systems contain probabilistic elements, they may be falsely interpreted as resulting from causal processes or mechanisms that do not exist. This is particularly problematic for coping with rare events and small numbers of cases because, for simple statistical reasons, they are more likely to produce large deviations from the expected values or normal outcomes.⁷⁰

Formal models are particularly useful for generating or evaluating alternative explanations of a system’s behavior for three reasons. First, because they are externalized representations of their developer’s assumptions and logic, their use can ensure that explanations do not include internal contradictions in their logic. Second, because models allow for researchers to explore many alternative configurations or sequences, they can formalize what analysts do via intuition—search across counterfactual cases until they have identified a set of circumstances that generate an outcome of interest.⁷¹ Formal models allow for the inclusion, removal, rearrangement, and adjustment of variables and actors within a system in order to facilitate the development of explanations of how the system works. For example, analysts seeking to

⁶⁹ Daniel Kahneman, *Thinking Fast and Slow* (New York, NY: Farrar, Strauss, and Giroux, 2011), pp. 203-204.

⁷⁰ Daniel Kahneman, *Thinking Fast and Slow* (New York, NY: Farrar, Strauss, and Giroux, 2011), pp. 109-111.

⁷¹ Philip E. Tetlock and Aaron Belkin, “Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives,” in Philip E. Tetlock and Aaron Belkin, eds., *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives* (Princeton, NJ: Princeton University Press, 1996), pp. 3-38.

explain Iranian nuclear energy policy may develop a model of the regime's decision making, and then experiment with alternative configurations of the regime until they can identify sets of conditions that produce results consistent with the real-world behavior they seek to explain.

Finally, formal models can offer null cases that can serve as a hedge against becoming overly invested in particular explanations by revealing alternatives. Because the unassisted human brain cannot deal with complexity in a rigorous fashion, errors in determining what evidence is consistent with what explanation are likely to occur, meaning that accepted explanations may be far from the truth. This links explanation with exploration, discussed later.

It is often assumed that explanation and prediction are two sides of the same coin, i.e., success at one naturally leads to success of the other. However, a recent review of government modeling efforts' failure to predict the Arab Spring noted that even in cases of successful prediction of world events, there is little understanding of why they occur and how they can be explained.⁷² This is because prediction and explanation are asymmetric, and constitute different forms of knowledge that do not translate easily. Because many predictive models are statistical in nature, they do not provide a causal explanation as to why a particular outcome is expected, even though they can identify correlations between observable, or inferred, variables.

⁷² Noah Shachtman, "Pentagon's Prediction Software Didn't Spot Egypt Unrest," *Danger Room, Wired* (February 11, 2011), <http://www.wired.com/dangerroom/2011/02/pentagon-predict-egypt-unrest/> (accessed on August 6, 2011).

Many of the emerging predictive techniques encounter this problem, as is the case with prediction markets that aggregate the perspectives of traders who bet on the likelihood of particular future events occurring, e.g. the likelihood of a foreign leader remaining in power or a percentage change in the price of oil.⁷³ In such cases, market prices and price movements may offer a prediction, but each trader's rationale for entering or leaving the market, buying shares in a future event occurring, or selling shares of that event, may be idiosyncratic and not explicitly revealed in the price of the prospective event. Therefore, determining the likelihood of the event is not the same as explaining why it might or might not occur. For example, the market may believe that a foreign leader is likely to lose power, but there is no way to know whether that is because traders believe that the leader will lose an election, leave office due to poor health, or be the victim of a coup. While new event scenarios can be introduced, specifying a future with greater precision, there remains no basis for converting the single dimension of price, or likelihood of occurrence, into a multidimensional narrative that explains the causal logic behind expectations. For this reason, several intelligence professionals responsible for the development and teaching of analytic tradecraft have questioned the utility of prediction markets precisely because they offer prediction without explanation.

Prediction Markets do have a strong record of near-term forecasts, but intelligence analysts and their customers are likely to be uncomfortable with their predictions. No matter what the statistical record of accuracy with this technique might be, consumers of intelligence are unlikely to

⁷³ James Surowiecki, *The Wisdom of Crowds* (New York, NY: Doubleday, 2004); and Robert E. Looney, "DARPA's Policy Analysis Market for Intelligence: Outside the Box or Off the Wall?" *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), pp. 405-419.

accept any forecast without understanding the rationale for the forecast and the qualifications of those who voted on it.⁷⁴

A second reason why explanations do not necessarily lead to predictions is that they are asymmetric. The knowledge that can predict the physical properties of an artifact does not explain the social, economic, organizational, or psychological reasons of its design.⁷⁵ For example, an imagery analyst may be able to predict and explain the length of a shadow cast by a foreign missile of known height by referring to geometry and the physics of light. However, the inverse is not true: Knowledge of a shadow's length can be used to predict the height of the missile, but it cannot explain why the missile was designed to be a particular height. Instead, such explanations reside in complex social and engineering issues such as military strategy, availability of particular technologies, economic costs, etc.

Given the existence of non-explanatory predictions and asymmetries between explanations, model users should be careful when using models. Users should not assume that models that predict a phenomenon necessarily explain it, or that explanations can always be projected forward in time in order to provide predictions.

Exploration

Exploration is a well-established use of models, particularly simulations and games, because of their dynamic properties. The earliest applications of computer simulation and modeling evolved at the junction between the mathematical modeling of

⁷⁴ Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), p. 210.

⁷⁵ Samir Okasha, *Philosophy of Science: A Very Short Introduction* (New York, NY: Oxford University Press, 2002), pp. 45-46.

military operations and the performance of wargames that sought to illuminate the competitive dynamics of future and potential armed conflict.⁷⁶ While early simulations aspired to predict the outcomes of potential scenarios, it became evident to model developers and users that such aspirations would always fall short for a variety of scientific, technical, and cultural reasons. Nevertheless, repeated surveys of model users and wargame players noted that participation in games enriched their understanding of the problems they studied and found that it allowed them to identify and challenge critical assumptions and reframe problems. As Brewer and Shubik noted, the use of operational games, i.e. those played to evaluate specific weapon systems or operational concepts, made their greatest contributions by stimulating new questions and refining research agendas in order to address the shortcomings or gaps in knowledge they expose.

A commonly overlooked use of operational games—the heuristic—is associated with their capacity to stimulate player, albeit constructive, explorations of problems that are either not well understood or misunderstood. Especially in free-form and scenario-based MSGs, the discovery or realization of unimagined difficulties and opportunities is a standard outcome. Using games in this manner may help to inform empirical inquiry, but it is by itself insufficient to achieve scientific understanding; follow-up investigations and studies must ascertain the validity and reliability of the heuristic insight.⁷⁷

⁷⁶ Andrew Wilson, *The Bomb and the Computer: A Critical History of War Games* (New York, NY: Delacorte Press, 1969).

⁷⁷ Andrew Wilson, *The Bomb and the Computer: A Critical History of War Games* (New York, NY: Delacorte Press, 1969); and Gary D. Brewer and Martin Shubik, *The War Game: A Critique of Military Problem Solving* (Cambridge, MA: Harvard University Press, 1979).

These techniques, pioneered by the military, were exported to industry and applied in a less formal fashion to problems of business strategy and uncertainty, most frequently through the use of scenario planning.⁷⁸

Exploration is a particularly important use of models in the social sciences, policy analysis, and intelligence. Model-based exploration provides analysts with opportunities to test and develop their intuition about pressing problems, and enables them to generate and explore alternative futures that could occur under different model configurations or even repeated samples of the same stochastic process. Models enable analysts to explore the limits of their imagination, i.e. alternative theories about the system, its design, and parameters, in order to look beyond what can be observed and gathered empirically in order to support the most difficult decisions that their consumers face—choosing between alternative options whose costs and benefits cannot be known *a priori*. Indeed, the inability to know beforehand whether a model can provide reliable or accurate assessments of the consequences of alternative choices has been a neglected epistemological problem in operations research since its founding during World War II. Despite claims of precision and prediction by modelers, Walter Strauss noted that the entire field suffered from an underappreciated and often neglected epistemological foundation that resisted empirical validation.

It appears to me that in general structure our epistemology, the process whereby our knowledge is obtained, is *similar* to that of a mathematical deductive system. We think we are able to nail down more firmly some rather elementary types of assumptions, or inputs into parts of the model,

⁷⁸ Peter Schwartz, *The Art of the Long View: Planning for the Future in an Uncertain World* (New York, NY: Doubleday, 1996); Kees van der Heijden, *Scenarios: The Art of Strategic Conversation* (New York, NY: John Wiley & Sons, Ltd., 2009); and Thomas Chermack, *Scenario Planning in Organizations: How to Create, Use, and Assess Scenarios* (San Francisco, CA: Berett-Koehler, 2011).

than the general statements that we try to deduce from these. Our faith in the reasonableness of these elementary assumptions is more deeply rooted than our faith in the general answers that we are trying to get.

Our epistemology is similar to that of the empirical sciences in that if we cannot accept the various statements deduced from our initial assumptions, we reserve the right to change the assumptions. Most people...agree that one part of the scientific method applied in the physical sciences, say specifically in physics, is that of testing hypotheses.... The predictiveness of a statement in physics can be checked by confronting it with a real-world situation. It should be stressed that our conclusions cannot be tested even “in principle.” Most of the weapon systems of concern to us do not exist. The testing of weapon systems, under conditions for which we want to study them, would presuppose the decision that we are trying to influence what had already been made. No control situation could exist. To put it tritely, one cannot eat one’s cake and have it too. Nor can one test recommendations made with respect to earlier decisions that have since been implemented. It would require an actual war to test such earlier recommendations. Even if this were possible, it would not test recommendations—since the choice was usually between competing systems or competing force compositions, only one of which was procured. Thus the comparison could not be made. Further, the recommendations were based on many different possible contingencies, only a few of which would actually be realized with the passage of time. This argues that the particular recommendations made, to affect a particular decision to be faced, cannot be tested. Can one then learn from tests carried out with other weapon systems, under other situations of the past or present world? The answer, of course, is, yes. To an extent, properties of the interaction of future weapon systems with future situations do exist in the interactions of present weapon systems with present situations. For example, one can examine the tracking process of present-day radars against present-day bombers. The testing of present-day equipment in present-day situations, such as in specifically planned exercises, is necessary to establish base points in our type of studies.⁷⁹

Given the character of the questions that intelligence producers and consumers face, deep consideration of how models are designed and used is essential to ensure that their epistemological limitations are constantly considered and that results are properly

⁷⁹ Walter J. Strauss, “The Nature and Validity of Operations Research Studies, With Emphasis on Force Composition,” *Operations Research*, Vol. 8, No. 5 (September/October 1960), pp. 685-686.

characterized and employed in decision-making processes. As a result, opportunities for prediction may be increasingly rare and limited, and explanations may be fleeting and limited, providing insights into sufficiency but not necessity regarding potential events. Meanwhile, exploration is likely the most effective and intellectually justifiable application of models because its primary purpose is to assist analysts and policy makers in exploring the consequences of their ideas, rather than make directly testable statements about the real world.

Partial Analysis

Partial analysis is a subtle but important use of models in analysis. Because models are simplifications, they provide a partial but incomplete insight into the system's behavior or properties—what they exclude is often as important as what they include. In cases where models are known to capture some aspects of a system, but not others, the model's results are used to provide a benchmark against which additional considerations create deviations. This usually occurs when the model being studied represents an idealization of a system, such as a political system in which all of the stakeholders are perfectly rational and have no uncertainties about what others might do. For example, Ben-Israel noted that rational models of human behavior were approximations of real-world systems that served to establish hypotheses about the target system, and did not constitute a useful model of a real-world target in isolation.

Intelligence problems are certainly complex. But this is true of all real, concrete problems, including those of physics. Is it possible to “solve” (that is, predict) weather (“clouds”)? Or the outcome of a fire in a wood? Concrete physical problems are also highly complex, and usually too

complicated to solve. In solving such technological problems, therefore, we have to use partial solutions or *approximations*.

An approximation is not, of course, a truth. On the contrary, the very use of the term indicates missing factors. From a logical point of view, then, an approximation is always false. But it is a *useful falsehood* for technology. Its user knows the approximate solution is not true, but merely gives him some insight into the true solution.

The use of approximation to overcome the complexity of real problems is as effective in intelligence as in natural sciences. Intelligence and social science in general, in fact, have an inherent advantage here: We can always get the first approximation assuming that the object of research (“enemy:”) behaves *rationally*. Confronting a problem with a large “cloudy” component, we usually need to know what decisions were made by certain individuals or societies. “Will Syria go to war next summer?” It depends on decisions taken by Syria’s leaders. “What will happen in Syria after President Assad?” The answer depends on the behavior of a whole society, or at least a large part of it. We can get some insight into the answer to these questions (“a first approximation”) by assuming that Assad, his partners, or all Syrian society behave rationally. That is, that they try to reach their goals (theirs, not ours) in an optimal way (as they, not we, see it).

It is clear that this will not give us a final “true” solution, free of any doubt. Man has an element of free will, and can always act irrationally, even against his own interests. But no approximation, even one in physics, is “true.” *The rationality of our enemy leaders (their rationality, not ours) is nothing but a hypothesis....*

Like any other approximation, “solutions” derived with the help of rationality hypothesis must be handled very carefully. Any action based on predictions derived from *rationality hypothesis* must be carried out piecemeal, under continuous control to detect the first deviation from what was predicted. This is what happens in technology. Construction of a new aircraft, for example, while based on scientific principles, is too complex for a full solution (like any concrete problem). Aeronautical engineers, therefore, use approximations, and it would be foolish to build an aircraft and fly it without first checking very carefully, bit by bit, that the machine really behaves as expected.⁸⁰

Partial solutions derived from models support an evolutionary or adaptive approach to problem solving and analysis. The use of models in partial analysis encourages adaptive

⁸⁰ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 693-694.

approaches to problem solving where each modeling result constitutes a single step on the way to making sense of a more complex system. After each step, adaptations to the existing model are made, new models are introduced, or problems are reframed in order to evolve a comprehensive understanding of a system through a series of analytical cycles rather than as a single comprehensive assessment.

Robert Keohane's *After Hegemony: Cooperation and Discord in the World Political Economy* provided a demonstration of how models can be employed for partial analysis. Keohane's main argument, made during the Cold War's final decade, was that the international regimes put in place to preserve the stability of the international system during the Cold War had the potential to persist and expand after the decline of great powers.⁸¹ In making his argument, Keohane's initial analysis of the logic of regimes was made within the context of rational states that possessed no limitations on their information gathering or processing capabilities, and rested on judgments based on formal models of market failure and economic institutions.⁸² After establishing a baseline based on an idealized representation of state behavior, Keohane advanced his argument with an informal model based on the theory of boundedly rational national leaders and governments. By arguing that because regimes helped perfectly rational actors cope with the costs of collective bargaining and uncertainty regarding each other's intentions and reliability, these same regimes would provide even greater benefit to

⁸¹ Robert O. Keohane, *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton, NJ: Princeton University Press, 2005).

⁸² Robert O. Keohane, *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton, NJ: Princeton University Press, 2005), pp. 65-109.

boundedly rational actors with more realistic cognitive and organizational information-gathering and processing capabilities.

Thus if we accept that governments must adopt rules of thumb, the costs of adhering to international regimes appear less severe than they would be if classical rationality were a realistic possibility. Regimes merely substitute multilateral rules (presumably somewhat less congenial per se) for unilateral ones, with the advantage that other actors' behavior thereby becomes more predictably cooperative. International regimes neither enforce hierarchical rules on governments nor substitute their own rules for autonomous calculation; instead, they provide rules of thumb in place of those that governments would otherwise adopt.

Combining this argument with that of [the functional theory of regimes based on rational actors and political market failure], we can see how different our conception of international regimes is from the self-help system that is often taken as revealing the essence of international politics. In a pure self-help system, each actor calculates its interests on each particular issue, preserving its options until that decision has been made. The rational response to another actor's distress in such a system is to take advantage of it by driving a hard bargain, demanding as much as "the traffic will bear" in return for one's money, one's oil, or one's military support. Many such bargains are in fact struck in world politics, especially among adversaries; but one of the key features of international regimes is that they limit the ability of countries in a particularly strong bargaining position (however transitory) to take advantage of that situation. This limitation, as we have stressed, is not the result of altruism but of the fact that joining a regime changes calculations of long-run self-interest. To a government that values its ability to make future agreements, reputation is a crucial resource; and the most important aspect of an actor's reputation in world politics is the belief of others that it will keep its future commitments even when a particular situation, myopically viewed, makes it appear disadvantageous to do so. Thus even classically rational governments will sometimes join regimes and comply with their rules. To a government seeking to economize on decisionmaking costs, the regime is also valuable for providing rules of thumb; discarding it would require establishing a new set of rules to guide one's bureaucracy. The convenience of rules of thumb combines with the superiority of long-run calculations of self-interest over myopic ones to reinforce adherence to rules by egoistic governments, particularly when they labor under the constraints of bounded rationality.⁸³

⁸³ Robert O. Keohane, *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton, NJ: Princeton University Press, 2005), pp. 115-116.

Keohane's analysis showed how models could be helpful by providing platforms to build upon within a larger argument, without needing to be comprehensive in their own right. By drawing upon formal models and the assumption of rationality, he established a benchmark that was subsequently extended in order to address details missing from the formal model.

Organization

Organization employs models to track and arrange data. Visual models are usually employed for these purposes, providing users with a template for arranging vast quantities of data and retrieving information quickly. Many SATs and other techniques, such as Strengths, Weaknesses, Opportunities, and Threats (SWOT) matrices, or Alternative Competing Hypotheses (ACH), used by intelligence analysts provide value by arranging data in ways that allow for cross-referencing of dependencies in systems or identifying gaps in available information.⁸⁴

Collaboration

Collaboration can often be facilitated by the use of models. Model development requires identifying critical assumptions and frameworks for studying a research problem or intelligence issue. While this process is often regarded as a benefit to the individual model developer, models may also be developed and used by teams working together to integrate their mental models into a single formal model or set of alternatives.

Sometimes teams may consist of methodologists with technical modeling skills and

⁸⁴ Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011).

analysts possessing deep substantive expertise on particular topics. Other times, analysts may possess the necessary modeling skills but may choose to develop and use a model as a group. In each case, the process of developing and using formal models externalizes how analysts think about intelligence issues and allows for colleagues to provide inputs into the model, either in the form of suggesting alternative frameworks, challenging assumptions or specific parameter settings, or recommending particular experiments in order to explore alternative scenarios or test hypotheses. Often, the benefits of collaboration have served as an implicit goal of modeling projects, and provide great value to analysts regardless of the quality or effectiveness of the resulting model.

The collaborative benefits of model use have been identified as a benefit of other SATs because they depersonalize arguments between analysts by shifting the source of disagreements from individuals to the operations of an external object that can be examined from multiple perspectives.⁸⁵ Advanced uses of models in policy analysis and decision support have begun to employ ensembles of alternative, competing models in order to identify how differences in perspectives on a system can affect expectations about its current and future state. These collaborative approaches challenge the traditional notion of finding optimal solutions to pressing policy problems, and seek to alleviate the intense bureaucratic pressures to control the design of models and input data in order to shape the design of a singular best course of action and allocation of resources and responsibilities across organizations. The use of ensembles of multiple models emphasizes the identification of alternative perspectives and trade-offs, and seeks to

⁸⁵ Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), pp. 3-8.

develop robust, adaptive solutions to complex problems. As a result, analysis improves as the number of alternative models and datasets increase—resulting in shifting bureaucratic incentives for collaboration between groups that would have competed under more traditional paradigms of policy analysis and decision making.⁸⁶

Illustration

Illustration employs models to demonstrate a product that will be developed in the future. An outline of a report, an architectural drawing, or an artist's rendering of a future weapon system all illustrate of future products and allow for a discourse between the projects' stakeholders and developers. In these cases, models are not necessarily part of a process of research and discovery, but depict results within a larger context of organizational decision making and resource allocation.

Management

Management is another use of models wherein organizations use them to allocate resources and set priorities. Managerial models are often employed to automate repetitive tasks, or search for particular patterns in data that indicate when an event warrants analysts' or decision makers' attention. Examples of managerial activities informed by models include systems that process data and notify analysts of instances of tax return filings that match patterns of cheating, or irregular international travel patterns that indicate potential terrorist or criminal activity.

⁸⁶ Robert J. Lempert, Steven W. Popper, and Steven C. Bankes, *Shaping the Next One Hundred Years: New Methods for Long-Term Quantitative Policy Analysis* (Santa Monica, CA: RAND, 2003); and Aaron B. Frank, "Pre-Conflict Management Tools: Winning the Peace," *Defense & Technology Paper, No. 11* (Washington, DC: Center for Technology and National Security Policy, National Defense University, February 2005).

Training

Training is another use of models. Rather than employ models to predict future outcomes or explore the unknown, training models are used to provide users with a specific experience for teaching purposes. Simulations and games have been important applications of models for these purposes. Visual diagrams depicting proper tactics and wargames have been used for centuries in military training, and flight simulators have played an essential role in training military and commercial pilots since the invention of the airplane. In fact, the use of models for training purposes has largely improved the job performance of many disciplines by allowing people to experience situations that are too difficult or dangerous to replicate in the real world, such as flying an aircraft in severe weather conditions or when experiencing a mechanical failure.⁸⁷

While the most popular uses of models in training are for tactical applications and learning particular tasks and skills, simulations and games have also been employed to train policy makers at the highest levels of government. One such application has been demonstrated by the Strategic Economic Needs and Security Exercise (SENSE) developed by the Institute for Defense Analyses and currently employed by the United States Institute of Peace. SENSE was originally developed to assist senior policy makers in post-conflict Bosnia, and more recently Poland and Iraq, manage the complex tasks of economic reconstruction and social and political reconciliation. Through role-playing, negotiation, and computational simulation, game players can learn about the extended consequences of their choices through a combination of interactions with stakeholders at

⁸⁷ Albert Helfrick, *Electronics in the Evolution of Flight* (College Station, TX: Texas A & M Press, 2004), pp. 154-157.

all levels of society via role-playing and propagating feedbacks throughout the system determined by computational models of the economy.⁸⁸ Other policy makers have noted the importance of simulations and games in their preparation for making major policy decisions. In these instances, models may not teach particular substantive lessons, but point toward important decision-making heuristics, such as remaining flexible and adaptive by expecting the emergence of unintended consequences from their choices.⁸⁹

Agent-Based Modeling

ABM is a particular type of computational modeling technique. It possesses specific theoretical or ontological premises regarding the structure of systems, and is a research methodology that is enabled by modern computing technology. As a theory, ABMs assume that the structure of systems emerges from the interactions between their units, resulting in macroscopic properties that cannot be deduced by examining the individual units of the system in isolation. For example, water molecules may create a wet, slippery liquid, but the property of being wet cannot be attributed to any of the individual molecules. Instead, the property of “wetness” emerges from the interactions between molecules.⁹⁰ These macroscopic structures, often called *emergent properties*, can be found in many of the challenges intelligence analysts study, whether they are arms races, the decision-making processes of foreign regimes, the dynamics of mass protest movements and riots, or fluctuations in global financial markets. Indeed, former Deputy

⁸⁸ United States Institute for Peace, *Strategic Economic Needs and Security Exercise (SENSE)*, <http://www.usip.org/sense> (accessed on September 10, 2012).

⁸⁹ Interview with Leon Fuerth, National Defense University, September 21, 2001.

⁹⁰ Valeria Ahl and T. F. H. Allen, *Hierarchy Theory: A Vision, Vocabulary, and Epistemology* (New York, NY: Columbia University Press, 1996), p. 146.

Director of Central Intelligence Douglas MacEachin regarded the perpetual novelty of emergent threats as one of the primary challenges facing intelligence analysts, and a driver for the need to improve analytic tradecraft.⁹¹

The central premise motivating ABM is that no centralized controller directs the actions of individual actors within the system. Instead, actors collect and use what information is available to them, and take action according to their individual capabilities and objectives, which may include competing, cooperating, or coordinating with others. Because the actors in the system are endowed with some degree of autonomy, i.e. the ability to determine what course of action moves them closer to their goals, they are referred to as agents to denote agency.⁹² The result is a simulated social system or artificial society made of agents interacting with one another, each pursuing their individual or shared objectives.

As a methodology, ABM employs computer simulation to study how different representations of individual decision making and action interact to generate outcomes. By emphasizing how agents in models gather information and make choices, users can see how the detailed descriptions of micro-level actors can generate particular macro-level outcomes. Moreover, by focusing on the attributes and behaviors of individual units in the model, ABMs allow for a more intuitive and realistic representation of real-world people and organizations when compared with other formal modeling methods.

⁹¹ Douglas MacEachin, "Analysis and Estimates: Professional Practices in Intelligence Production," in Jennifer E. Sims and Burton Gerber, eds., *Transforming U.S. Intelligence* (Washington, DC: Georgetown University Press, 2005), p. 130.

⁹² Mark d'Inverno and Michael Luck, *Understanding Agent Systems* (New York, NY: Springer, 1998); and Michael Wooldridge, "Intelligent Agents," in Gerhard Weiss, ed., *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence* (Cambridge, MA: MIT Press, 1999), pp. 27-77.

Whereas many mathematical models make the assumption of perfect rationality, ABMs' emphasis on decision-making processes capitalizes on the detailed psychological and organizational descriptions of how real people and organizations make choices, allowing for more realistic simulation of their boundedly rational behavior.

As a technology, ABMs are computer programs that, like all software, are enabled and constrained by the ways in which computers work. Generally speaking, models can be developed that operate effectively within the bounds of available computing capabilities. However, there may be times in which the limitations of software programs, computer hardware, or even fundamental problems of computational complexity, may affect how a mental model or theory is formalized. Just as mathematical models often require particular assumptions about the system in order to create a usable model, computational models, including ABM, also require trade-offs between the best available description of a system and what can be practically implemented and experimented with. While the methodology of ABM allows for the relaxation of assumptions often found in mathematical models, the dependence on computers nevertheless requires the need for simplification in order to create usable software.

Origins of Agent-Based Modeling

The development of ABM as an analytic technique marks a convergence of different lines of research and disciplines. The concept of multi-agent systems grew out of the study of Artificial Intelligence (AI) in computer science as it transitioned from examining intelligent systems as single, highly complicated programs toward the concept of Distributed Artificial Intelligence (DAI) in which systems were collections of small,

simple programs that worked together to solve problems.⁹³ Simultaneously, social scientists such as Thomas Schelling, Herbert Simon, James March, and Harold Guetzkow sought to explore how cognitive and social processes affected decision making, strategy, and organizational behavior.⁹⁴ ABM marks the convergence of these research approaches.

Although social scientists have employed computer simulations since the invention of the computer, the development of the personal computer encouraged the proliferation of computer-based research methods.⁹⁵ During the 1980s and '90s, computer scientists, mathematicians, biologists, and social scientists engaged in a series of interdisciplinary research projects that developed the agent-based research paradigm, merging the technical foundation of AI with social theories of decision making and collective behavior. These projects varied from examining the movements of flocks of birds and schools of fish to the development of dynamic strategies of conflict and cooperation between nations.⁹⁶ By the end of the 1990s, ABM had emerged as a term of art, and had entered into the methodological toolkit for exploring the behavior of complex

⁹³ Mark d'Inverno and Michael Luck, *Understanding Agent Systems* (New York, NY: Springer, 1998); and Gerhard Weiss, ed., *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence* (Cambridge, MA: MIT Press, 1999).

⁹⁴ Michael Don Ward, ed., *Theories, Models, and Simulations in International Relations: Essays in Honor of Harold Guetzkow* (Boulder, CO: Westview Press, 1985); James G. March, *Decisions and Organizations* (New York, NY: Basic Books, 1988); Herbert A. Simon, *The Sciences of the Artificial* (Cambridge, MA: MIT Press, 1996); and Thomas C. Schelling, *Micromotives and Macrobehavior*, (New York, NY: W.W. Norton & Company, 2006).

⁹⁵ Irving Louis Horowitz, ed., *The Rise and Fall of Project Camelot* (Cambridge, MA: MIT Press, 1976); Nigel Gilbert and Rosaria Conte, eds., *Artificial Societies: the Computer Simulation of Social Life* (New York, NY: Routledge, 1995); and Orville Vernon Burton, ed., *Computing in the Social Sciences and Humanities* (Urbana, IL: University of Illinois Press, 2002).

⁹⁶ For examples see Craig Reynolds, *Boids: Background and Update*, <http://www.red3d.com/cwr/boids/> (accessed on August 7, 2011); and Robert Axelrod, *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration* (Princeton, NJ: Princeton University Press, 1997).

adaptive systems, particularly in economics, political science, and ecology.⁹⁷ At this time, computer simulations were increasingly recognized as a laboratory for performing experiments *in silico*, allowing for the manipulation of artificial societies in ways that cannot be performed in the real world.⁹⁸ As R. Keith Sawyer described:

The research can use these simulations to create *artificial societies* and to run “virtual experiments”—in which properties of agents...are varied and the subsequent change in the overall macrobehavior of the system are observed. Multi-agent systems have been used by complexity researchers to simulate a wide range of natural systems, including sand piles, industrial processes, and neuronal connections in the human brain; in the late 1990s, this methodology was increasingly used to simulate social systems.⁹⁹

The experimental opportunities that ABM provides have allowed for researchers to shift their attention to three questions of great importance to intelligence analysts and policy makers.

First, the ability to simulate a social environment from a predefined initial condition has enabled researchers to explore and search across counterfactuals, examining how alternative decisions by individuals or groups of agents might have produced different outcomes. Second, the use of simulation-based models allows for researchers to shift their attention to the model’s dynamics rather than solving for

⁹⁷ Joshua M. Epstein and Robert Axtell, *Growing Artificial Societies: Social Science from the Bottom Up* (Cambridge, MA: MIT Press 1996); W. Brian Arthur, Steven N. Durlaf, and David Lane, eds., *The Economy as an Evolving Complex System II* (Boulder, CO: Westview Press, 1997); Robert Axelrod, *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration* (Princeton, NJ: Princeton University Press, 1997); Lars-Erik Cederman, *Emergent Actors in World Politics: How States and Nations Develop and Dissolve* (Princeton, NJ: Princeton University Press, 1997); Simon Levin, *Fragile Dominion: Complexity and the Commons* (Cambridge, MA: Perseus Publishing, 1999); Volker Grimm and Patrick Railsback, *Individual-Based Modeling and Ecology* (Princeton, NJ: Princeton University Press, 2005), and Simon A. Levin, ed., *The Princeton Guide to Ecology* (Princeton, University Press, 2009).

⁹⁸ Joshua M. Epstein and Robert Axtell, *Growing Artificial Societies: Social Science from the Bottom Up* (Cambridge, MA: MIT Press 1996), pp. 52-53.

⁹⁹ R. Keith Sawyer, *Social Emergence: Societies as Complex Systems* (New York, NY: Cambridge University Press, 2005), p. 2.

equilibrium or steady-state outcomes. Third, by emphasizing the social interactions of autonomous agents, representations of systems focus on the identification and description of discrete actors. As a result, an agent-based approach to simulation allows for the decomposition of systems into parts that make intuitive sense to users, i.e. natural units, rather than rely on aggregated, representative units that depict the average person or organization—particularly when the actors of greatest attention might be outliers within the population, such as foreign elites or terrorists.¹⁰⁰ Moreover, the emphasis on disaggregation and interaction allows for the consideration of features of systems that are often beyond the ability of other models to address, such as the role of culture, morale, intelligence, and command and control in military organizations.¹⁰¹

By providing a means for counterfactual analysis and the study of dynamics resulting from interactions between diverse actors, ABM fills an important gap in the theoretical and methodological toolkit available to intelligence analysts. Indeed, in 1991 Stephen Peter Rosen noted the inadequacy of classical mathematical models for studying the future of warfare and the dynamics of strategic balances of power:

If a computer could be programmed with the operational characteristics and observable capabilities of the United States and the Soviet Union, simulations might be conducted that would reveal the character of future wars with somewhat greater reliability and reproducibility than is the case with other analytic methods.

War is a complex social and political activity. Developing a game or model that captures all of its relevant features is approximately as difficult as developing a model of human society. As a result, all war games and political-military simulations, whether they are conducted by computers or human players, have to date had to utilize grossly simplified

¹⁰⁰ Interview with Simon A. Levin, Princeton University, April 20, 2010.

¹⁰¹ Mark Herman, “Entropy-Based Warfare: Modeling the Revolution in Military Affairs,” *Joint Forces Quarterly*, No. 40 (Winter 1998-99), pp. 85-90.

models of behavior in order to make them playable in reasonable amounts of time. In the case of group combat, each soldier's performance cannot be modeled. Differences among actual army units would complicate the model, so hypothetical standard military units are assumed. In actual combat the bravery and competence of the commander and his troops, morale, surprise, deception, fatigue, disrupted communications, and the like introduce far more complexity than can easily be captured by a computer or by an individual manipulating game pieces on a game board.¹⁰²

By the end of the 1990s, ABM and computing technology had developed to the points that what Rosen deemed impractical at the start of the decade was possible by the conclusion—a point that he noted in later research on the origins and nature of warfare, when he suggested the potential benefits of ABM-based research on warfare.¹⁰³

Underspecification and Mathematical Models

The differences between ABMs and classical mathematical models are best demonstrated through direct comparison. While mathematical models are often touted for their precision, they are in fact often underspecified in many ways that interest social scientists, intelligence analysts, and policy makers. This is evident by examining the process by which mathematical models are agentized, or converted into ABM, which requires unpacking many assumptions about information, space, time, and decision making that are not explicitly revealed by mathematical equations.

Paul Krugman developed a simple mathematical model of how firms organize in space. He argued that the question of space, i.e. the location where economic activities

¹⁰² Stephen Peter Rosen, "Net Assessment as an Analytical Concept," in Andrew W. Marshall, J. J. Martin, and Henry S. Rowen, eds., *On Not Confusing Ourselves: Essays on National Security Strategy in Honor of Albert & Roberta Wohlstetter* (Boulder, CO: Westview Press, 1991), p. 297.

¹⁰³ Stephen Peter Rosen, *War and Human Nature* (Princeton, NJ: Princeton University Press, 2005), pp. 180-181.

such as production, exchange, and consumption occurred, was an undertheorized aspect of economics.¹⁰⁴ To examine the ways in which economic firms organized spatially, Krugman developed a relatively simple model that incorporated two opposing forces that firms considered. Calling this the Edge City model, Krugman noted that businesses simultaneously attract and repel one another in space. Thus, businesses largely prefer to be collocated, where each benefits from the customer base that the others attract, but will also disperse and position themselves where they are not competing for the same customers. This model is depicted by the equation below:

$$P(x) = \int_z [A \exp(-r_1 D_{xz}) - B \exp(-r_2 D_{xz})] \lambda(z) dz$$

- $P(x)$ = the market potential of location x
- A = the attractive force between firms
- B = the repulsive force between firms
- r_1 = the rate of decay of the attractive force over distance
- r_2 = the rate of decay of the repulsive force over distance
- D = the distance between site x and site z
- $\lambda(z)$ = the density of firms at location x

The dynamics of the Edge City model are governed by each site's relative fitness as compared with the average fitness of all sites on the landscape. Average fitness is denoted by the term \bar{P} . Thus, the changing density of businesses at location x obeys the equation below.

¹⁰⁴ Mashahisa Fujita, Paul Krugman, and Anthony J. Venables, *The Spatial Economy* (Cambridge, MA: MIT Press, 2001); Paul Krugman, *The Self-Organizing Economy* (Malden, MA: Blackwell, 1996), pp. 101-106; and Paul Krugman, "How the Economy Organizes Itself in Space: A Survey of New Economic Geography," in W. Brian Arthur, Steven N. Durlaf, and David A. Lane, eds., *The Economy as a Complex Adaptive System II* (Boulder, CO: Westview Press, 1997), pp. 239-262.

$$\frac{d\lambda(x)}{dt} = \gamma[P(x) - \bar{P}]\lambda(x)$$

- $d\lambda(x)$ = change in the business density at location x
- dt = change in time
- γ = a constant that affects the rate of movement
- $P(x)$ = the market potential of site x
- \bar{P} = the average market potential of all sites
- $\lambda(x)$ = the density of firms at location x

The result of the model is that for a given arrangement of parameters, the result should produce equally spaced clusters of economic activity, where many businesses collocate on some locations while leaving others empty. Thus, a typical result starting from a random distribution of businesses would result in the agglomeration of businesses on a few locations as shown in Figure 2-3.

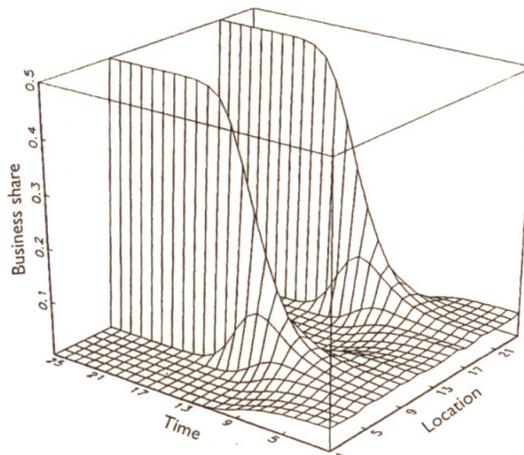


Figure 2-3: Emergence of Spatial Clusters in the Edge City Model. The figure shows how the random distribution of businesses cluster over time and produce two distinct centers of economic activity. In the figure, the initial location of each business is shown on the right axis, while each time step is shown on the left axis where time increases from the right to the left of the graph. The vertical axis depicts the percentage of businesses at each location, and shows the emergence of two fins that each contains 50% of the population at the simulation's end. Image from Paul Krugman, *The Self-Organizing Economy* (Malden, MA: Blackwell, 1996), p. 25.

Despite the clarity of the Edge City model, it is underspecified from the agent-based perspective. First, the Edge City model makes no mention of how businesses understand their environment and choose to move. For example, do businesses survey the entire landscape and then move to a new location based on calculating the best site they identify? Do they examine nearby locations and move to the best available site within a limited search distance? Do they simply move to the closest location whose business potential is higher than their current location, even if it is not the optimal location available to them? From the agent perspective, the behavioral processes that govern agents' decision making is not specified in the equation, but different processes would produce different transient dynamics, and potentially reach different equilibrium (or even no equilibrium) than the mathematical representation.

Secondly, the Edge City model does not explain how businesses move with respect to time. Do all agents move simultaneously, or does one move first and then another and another and so on? Does the location of other agents change between the time they evaluate the market potential of different sites and actually move?

Finally, while the Edge City model purports to study the movements of businesses, its actual interpretation is far more difficult. The representation of business as continuous, increasingly divisible units that make smooth spatial transitions is problematic. For example, the conclusion that 50% of the businesses can be found at two locations cannot be achieved in a system with an odd number of agents. Once businesses are seen as heterogeneous, discrete actors of different size, the model is increasingly difficult to interpret as to how firms locate in space. Instead, the mathematical model

makes a stronger case for describing the movement of market share: depicting where consumers go to spend their money.¹⁰⁵ This would then mean that businesses are not relocating but that consumers flock to some locations and avoid others. As a result, a more precise interpretation of the model may be that it characterizes what agents and firms will go out of business and which will thrive based on their location and the movement of market share.

By examining the Edge City model from the perspective of the units within the model, agents as businesses located in space, the problems that ABM seeks to address become readily apparent. The heterogeneous behavior and size of the businesses, the ways in which they interact and the information at their disposal, and other features are all underspecified when modeling mathematically, yet serve as the foundation of an agent-based approach to thinking about models. In this sense, ABMs seek to address the features in which mathematical models are imprecise and poorly communicated by providing alternative ways of representing systems in ways that are consistent with the ontologies and behavioral descriptions found in the verbal models of the social sciences and intelligence analyses.

The Benefits of Agent-Based Modeling

ABMs provide analysts with unique opportunities to study complex intelligence problems. ABMs vary in innumerable ways, covering the gamut from simplicity to great complexity with respect to the quantities of agents, diversity of agent types,

¹⁰⁵ Agent-Lab, *Evolution of Central Places (Urban Morphogenesis)*, http://www.maml.hu/maml/model_summaries/EdgeCity/index.html (accessed on September 10, 2012).

sophistication of their logic, range of possible interactions, and environments. Yet, despite the flexibility of the methodology and technology, models share a core structure that focuses on how agents interact with one another and their environments in order to generate emergent structure and properties from the bottom up.

ABMs enable users to overcome cognitive limitations with respect to complex adaptive systems. This ability provides many new opportunities to analyze challenges that involve large-scale, decentralized interactions, such as the spontaneous emergence of riots or segregated communities, as well as animate the long-run trajectories of individual and group decision-making processes. They also allow for representations of systems in ways that are more intuitive than equations, because they maintain the discrete character of the units within the model. This enables ABMs to arrive at conclusions that could not be predicted using more conventional mathematical models, while providing a self-contained artificial society against which experiments can be performed.

Chapter 3: Intelligence Analysis and Tradecraft

Intelligence analysis has undergone a significant transformation over the last two decades. Even before the events of 9/11 and the 2002 National Intelligence Estimate regarding Iraq's weapons of mass destruction (WMD) capabilities, the intelligence community had been working toward improving analytic tradecraft by changing the character of analytic production from one in which analysts primarily saw their value as "making the call" for consumers, to ensuring that policy makers understood the basis of critical uncertainties about the international system, the implications they posed with respect to analytic judgments, and the body of evidence and logic upon which inferences were made.

This chapter examines the institutional, cultural, and methodological problems of intelligence analysis and analytic tradecraft. Intelligence analysis differs from traditional scholarship or journalism because of the institutional context in which analysts operate. As a result, intelligence problems cannot be understood without reference to the policy-making context in which consumers consider, accept, or reject analytic products. Thus, analysts face the normal difficulties of trying to understand the behavior of intelligence targets while ensuring that their own biases, preconceptions, experience, and other cognitive limits do not affect their judgments, while also seeking to ensure their assessments will be considered timely, accurate, and relevant by consumers. Therefore,

this section begins with an examination of the general conceptual challenges analysts confront and that tradecraft must address, as well as an overview of the significant institutional, cultural, and epistemological issues that complicate relations between producers and consumers.

After discussing producer/consumer relations, the specific development and uses of SATs are examined. SATs constitute the basis of contemporary analytic tradecraft, and ground analytic practices in the problem of bounded rationality discussed in Chapter 2. This discussion includes the addition of cognitive processes and limitations into the study of intelligence failure, and the transition from discussions of mindsets to mental models as a framework for thinking about analytic practices, processes, and expertise. In this examination, the general features of SATs and how they fit within contemporary tradecraft are discussed, as well as the application of two prominent methods, the use of alternative futures (scenarios) and Alternative Competing Hypotheses (ACH). The chapter concludes by noting the strengths and weaknesses of SATs as they are currently practiced, and notes that they constitute an important, transitory step toward the use of formal models in intelligence analysis, setting the stage for the scientific practice of intelligence analysis discussed in the following chapter.

The Three-Level Problem of Intelligence Analysis

Intelligence analysis constitutes a three-level problem in which analysts must understand the capabilities, intentions, and circumstances of other actors who possess their own worldviews, concerns, and expectations. On one level, analysts are constantly challenged by agency, i.e. the freewill of the targets of the intelligence collection and

analysis, which demands that analysts give due weight to the ways in which their subjects perceive, think, and behave. On another level, intelligence analysts must confront their own individual and collective psychology and cognitive limitations—a necessary addition to their technical or subject matter expertise. Finally, analysts must understand how consumers think. This means that analysts must not only strive to understand the world through the perspectives of intelligence targets and account for their own mindsets, but they must be capable of translating their assessments into a language that is meaningful to consumers.

Understanding the behavior and actions of intelligence targets is no different than the work of social scientists or journalists who seek to characterize and describe the behavior of their subjects. This is not a straightforward activity since targets may engaged in deceptive practices, seeking to deny or manipulate intelligence collection, resulting in most descriptions of their motivations to rest on inferences made from fragmented, ambiguous, and potentially deceptive evidence.

Given that few intelligence analysts have unfettered access to the private thoughts of their targets, the drivers of their analysis are their own beliefs about the target, or rather their beliefs about what the target believes.¹⁰⁶ As a result, intelligence analysts must give equal weight to their own beliefs and cognitive process that serve as the lens through which information about the target is perceived. Bruce and George differentiated intelligence analysts from other subject matter experts based on the requirement to engage in self-critical assessments of their own perceptions and beliefs.

¹⁰⁶ Robert Jervis, *The Logic of Images in International Relations* (New York, NY: Columbia University Press, 1970).

Many so-called subject matter experts are well versed in the history, politics, culture, and language of many countries or are technical experts in a wide variety of areas; they may also be very attuned to U.S. policy deliberations and indeed be involved in advising a number of government officials on the correct policies to adopt. And many foreign affairs specialists may have methodological expertise. Where the intelligence analyst distinguishes himself or herself is in having the other four characteristics [understanding unique intelligence collection sources and methods, self-awareness of cognitive biases and influences, open-mindedness to contrary views and alternative models, and self-confidence to admit analytic errors]. The complete analyst must be an expert on how to use intelligence collection capabilities; be both imaginative and rigorous in considering explanations for missing, confusing, and often contradictory data while at the same time being able to be a self-critic of one's own biases and expectations of what the data show; and, most important, be open to changing one's mind and consciously trying to ask the question, "If I'm wrong, how might I need to modify the way I am analyzing the problem?"¹⁰⁷

Finally, analysts must understand how consumers think. If consumers are not receptive to intelligence products then the entire effort may be for naught. This means that analysts must not only strive to understand the world through the perspectives of intelligence targets and account for their own mindsets, but they must be capable of translating their assessments into language that is meaningful to consumers. This point was noted by Gregory Treverton, who discussed how intelligence analysts were unable to communicate North Vietnamese resolve in the face of the American strategic bombing campaign to Lyndon Johnson due to a lack of understanding Johnson's framing of the problem.

Yet Johnson also seems to have had in his head another analogy, this one of North Korean leader Ho Chi Minh, one that likened him to a recalcitrant U.S. senator. If Johnson could find just the right combination of carrots and sticks, expressed with the right combination of bluster and

¹⁰⁷ James B. Bruce and Roger Z. George, "Intelligence Analysis—The Emergence of a Discipline," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, Innovations* (Washington, DC: Georgetown University Press, 2008), p. 3.

flattery, surely Ho would see reason. And so the bombing campaigns were accompanied by extravagant promises of public works projects that the North might receive once it made peace. For intelligence to realize that Johnson carried that analogy would have taken serious thought about his personal history and his formative experiences as Senate majority leader. Once it realized the analogy, providing more detail on Ho's own history of struggle might have challenged it. A man who had spent his entire adult life fighting for Vietnamese nationalism was not likely to be much swayed by promises of dams and electrification.¹⁰⁸

These three challenges characterize what Mark Lowenthal characterized as “thinking about our thinking while we are thinking—a mental triple play.” Indeed, behind the obvious challenges of studying the behaviors of targets and systems in real time and projecting future possibilities lies the equal problem of matching analytic products, in both substance and style, to the needs of policy-making consumers. Thus, while the topics of intelligence are external, i.e. focused on the potential activities and interactions of intelligence targets in the international system, the intellectual challenges addressed by analytic tradecraft are internal and focused on exposing and challenging cognitive biases of analysts while providing meaningful assessments to policy makers.

Producer/Consumer Relations

The relationship between intelligence analysts who produce intelligence, and policy makers who use or consume intelligence as an input in their decision making, is one of the most difficult, persistent, but least-studied problems in the intelligence

¹⁰⁸ Gregory F. Treverton, *Reshaping Intelligence for an Age of Information* (New York, NY: Cambridge University Press, 2003), pp. 186-187.

community.^{109, 110} Since the inception of the modern intelligence community, Sherman Kent noted that producer/consumer relations were the most important aspect of all intelligence work, since the very purpose of the community's existence is to serve the needs of policy makers.

There is no phase of the intelligence business which is more important than the proper relationship between intelligence itself and the people who use its product. Oddly enough, this relationship, which one would expect to establish itself automatically, does not do this. It is established as a result of a great deal of persistent conscious effort, and is likely to disappear when the effort is relaxed.

Proper relationship between intelligence producers and consumers is one of utmost delicacy. Intelligence must be close enough to policy, plans, and operations to have the greatest amount of guidance, and must not be so close that it loses its objectivity and integrity of judgment.¹¹¹

Six decades later, John McLaughlin noted that Kent's concerns and insights into the difficulties of this relationship remained the central problem in the practice of intelligence analysis.

If [Kent's argument about the relationship between producers and consumers is true], as I believe it is, that analysis is where all aspects of the intelligence profession come together, then it is equally true that dealing with the policymaker is where all the components of analysis come together. It is at the nexus between intelligence and policy that we test everything from the substantive merit of the product to the quality of our tradecraft to our effectiveness in training and managing analysts. And it is also where an analytic profession that strives for objectivity, civility, thoroughness, and balance is likely to meet up with the more jarring qualities—urgency, impatience with nuance or equivocation, and, yes, sometimes even politics. But if this relationship turns sour—if the policymaker does not feel the need for the analytic product—then there is

¹⁰⁹ Intelligence producers may also include collectors who gather intelligence information and managers who allocate resources across the community and determine priorities for collectors and analysts. In this section, producers will primarily refer to analysts unless otherwise noted.

¹¹⁰ Joshua Rovner, *Fixing the Facts: National Security and the Politics of Intelligence* (Ithaca, NY: Cornell University Press, 2011), pp. 3-6.

¹¹¹ Sherman Kent, *Strategic Intelligence for American World Policy* (Princeton, NJ: Princeton University Press, 1949), p. 180.

no reason for doing analysis at all. It goes without saying, then, that it is worth thinking about what makes the relationship work and what renders it dysfunctional.¹¹²

The relationship between producers and consumers is complicated by three challenges. First, intelligence is simultaneously independent of, and subservient to, policy makers. Second, producers and consumers maintain different beliefs about agency and the strategic character of the international system; as a result, significant cultural tensions exist between the two. Finally, intelligence production maintains a particular epistemological structure that always allows producers to dismiss the work of analysts as theoretical or uncertain. As a result, the relationship between intelligence analysts and policy makers tends to be characterized by competition and tension more than cooperation and harmony.

Before discussing the known problems of producer/consumer relations and the ways in which ABMs might address them, it is important to discuss the characteristics of prediction in the international system from the perspective of policy making and intelligence. This is necessary because many of the criticisms of intelligence analysis from outside of the community erroneously focus on successful predictions about international events as the measure by which analytic products should be evaluated. However, this model is fundamentally problematic and must be addressed in order to clarify why many of the proposed remedies for improving intelligence analysis are misguided.

¹¹² John McLaughlin, "Serving the National Policymaker," in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 71.

Prediction and Idealization

One of the consequences of the limited study of producer/consumer relations has been the mischaracterization of intelligence analysis at institutional and philosophical levels of analysis. From an institutional perspective, harmonious relations between producers and consumers create the illusion of simple fixes to dilemmas of governance and policy making. In an idealized model of producer/consumer relations, intelligence analysts provide policy makers with predictions about the future and recommend courses of action that will achieve the policy makers' desired goals with the highest likelihood of success, minimal cost, and lowest risk. In this model, a straight line exists between intelligence analysis and the choices of policy makers: Better predictions facilitate better policy.

This model is problematic because it rests on a technocratic formula for displacing the judgments of elected and appointed officials with a professional class of analysts that may possess highly developed technical expertise but no ethical or institutional obligations to democratically or constitutionally proscribed sources of legitimacy. As a result, it strips policy makers of their institutional power and prerogatives, while elevating analysts to positions of authority via their ability to control resources via agenda setting and resource allocations based on their assessments. As Mark Lowenthal noted, the primacy of the policy maker is one of the foundational elements of producer/consumer relations.

First, the relationship is and should be dominated by the policymakers, who have contested and won an election. They have the right to govern, to make decisions, to create budgets and to order operations. Second, intelligence is a service that is provided to the policymakers. It is an

important and useful part of the policy process but its role is determined by the policymakers, not by the intelligence agencies.¹¹³

Ironically, policy makers have strong incentives for politicizing intelligence in order to shield themselves from accountability for their decisions. This behavior is captured in the popular saying that “There are no policy failures. There are only policy successes and intelligence failures.”¹¹⁴ By justifying policy on intelligence predictions alone, policy questions that require strategic and ethical judgment can be transformed into intelligence questions about collection and analysis. As Paul Pillar noted, this then politicizes intelligence by publically employing conclusions drawn from highly uncertain, secret sources, methods, and assessments of independent analysts, to justify policy choices on the grounds that no alternatives exist.

Policymakers have strong reasons to try to use intelligence in publicly selling their policies. Because intelligence is supposed to be objective, it bolsters the credibility of any sales campaign. It adds what are perceived as hard facts—from sources that skeptics may find difficult to question—to what might otherwise be dismissed as mere exhortation from policymakers. It can make an act of choice appear to be one of necessity. Intelligence adds authority to any case for a policy.¹¹⁵

Therefore, prediction carries with it deeply problematic implications for producer/consumer relations and governance in general. It militates against democratic governance by transferring power to a class of unelected technocrats, while alleviating

¹¹³ Mark M. Lowenthal, “The Policymaker-Intelligence Relationship,” in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), p. 450.

¹¹⁴ Martin Petersen, “What We Should Demand from Intelligence,” *National Security Studies Quarterly* Vol. 5, No. 2 (Spring 1999), p. 107. Also see Roger Z. George, “The Art of Strategy and Intelligence” in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 118; and Mark M. Lowenthal, “The Policymaker-Intelligence Relationship,” in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), p. 438.

¹¹⁵ Paul R. Pillar, “The Perils of Politicization,” in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), p. 474.

policy makers of the burden of judgment, giving them a free hand to take risks by shifting the consequences of their failures onto the intelligence community.

A secondary problem with the emphasis on prediction is that it promotes a vision of the international system and human affairs that is deterministic. Such a view is ironic, however, since a predictable world is necessarily where an *a priori* knowable path cannot be changed by policy makers' choices and actions, whether individually or collectively. The possibility of successful prediction and harmony in producer/consumer relations undermines the need for policy in the first place, rendering decision makers an unnecessary appendage to rational, scientific governance. Focusing on the accuracy of intelligence analysts' predictions recasts fundamental strategic questions about complex systems as simple problems of optimization and risk management—denoting a peculiar commitment to deterministic sciences that are increasingly viewed as special cases in the broader universe of scientific inquiry.¹¹⁶

The ideal model of producer/consumer relations belongs to the same class of models as that of total war, whose theoretical parsimony and logic was necessarily rejected after confronting the empirical record and historical experience of political and military leaders. In the case of abstract, pure models of warfare, Clausewitz noted that the theory of maximum effort—the logical conclusion of total war—failed to correspond to the limited exertion of force by states and the persistence of political calculations governing their use.¹¹⁷ Just as Clausewitz argued that a pure theory of war provided

¹¹⁶ The issue of determinism in science and philosophy is discussed in greater detail in Chapter 4.

¹¹⁷ Carl von Clausewitz, *On War* (Princeton, NJ: Princeton University Press, 1984), Michael Howard and Peter Paret, trans., p. 78.

deductions of extreme violence, viewing producer/consumer relations as an ideal relationship developed around a predictable international system characterizes a set of experiences entirely foreign to real-world intelligence analysts and policy makers. As James Steinberg noted, while the relationship between analysts and policy makers should be harmonious, it is in fact acrimonious and filled with mistrust and tension as both sides fail to meet the expectations of the other.

Policymakers crave good intelligence. Why? Because they believe it can and should make the crucial difference between success and failure, at both the policy and personal levels. This should be the recipe for a match made in heaven between the intelligence analyst and the policymaker. Yet the reality.... Analysts typically feel underappreciated, ignored, or misused by policymakers, while policymakers in turn often feel misled or underserved by intelligence.¹¹⁸

Given the institutional and philosophical problems noted above, any effort to improve intelligence analysis must first and foremost consider the relationship between producers and consumers, since the context of their interactions ultimately determines whether analytic products are accepted or rejected by the policy process.

Institutions, Culture, and Epistemology in Producer/Consumer Relations

If the predictive accuracy of intelligence assessments does not characterize the relationship between producers and consumers, what does? As the passages from Lowenthal and Pillar cited above suggest, the relationships between the two are determined by their respective institutional roles in the policy process, and the alternative cultures that each institution promotes. In this context, uncertainty, rather than

¹¹⁸ James B. Steinberg, "The Policymaker's Perspective: Transparency and Partnership," in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 82.

prediction, takes on a special meaning due to the underlying epistemological basis upon which assessments can be made and how alternative competing perspectives on the international systems and policy actions can be justified. What follows from the study of producer/consumer relations is a concern with a more complicated problem—the non-use of intelligence—where for reasons of institutional politics, culture, and epistemology intelligence assessments are ignored, rejected, or attacked by the very policy makers they seek to support.¹¹⁹

Institutional Challenges

The institutional arrangement between producers and consumers places intelligence analysts in a position that is both independent of, and subordinate to, policy makers. As a result, how closely analysts and policy makers should work together remains an unsolved and perpetual problem.¹²⁰ In institutional terms, this is often regarded as the problem of politicization, or more specifically the extent to which intelligence analysts and organizations can or should tailor their products to suit the needs and interests of policy makers. At one end of this spectrum rests intelligence that is produced without regard to policy makers' interests, issues, and concerns, which as Treverton noted, risks irrelevance.

¹¹⁹ Sherman Kent, "Estimates and Influence" *Studies in Intelligence*, Vol. 12, No. 3 (1968), pp. 11-21; Amos Kovacs, "The Nonuse of Intelligence," *International Journal of Intelligence and Counterintelligence*, Vol. 10, No. 4 (December 1997), pp. 383-417; and Richard K. Betts, "The New Politics of Intelligence: Will Reforms Work This Time?" *Foreign Affairs*, Vol. 83, No. 3 (May/June, 2004), p. 7.

¹²⁰ Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York, NY: Columbia University Press, 2007), pp. 66-103; Josh Kerbel and Anthony Olcott, "Synthesizing with Clients, Not Analyzing for Customers," *Studies in Intelligence*, Vol. 54, No. 4 (December 2010), pp. 11-27; and interview with Paul Pillar, Georgetown University, February 2, 2012.

...too often, intelligence products seemed to answer questions no one was remotely asking. The questions were either ones that interested the analysts or safe ones. Too often they amounted to “Whither China?” But the policy community virtually never asks “Whither China?” Perhaps it should, especially for the purpose of long-term planning, but that planning is as rare as hen’s teeth.

The questions that policy officials ask are usually specific, time sensitive, and operational. And sometimes intelligence analysis seems driven by no question at all. It is akin to all those newspaper op-eds we read, which leave us scratching our heads and asking ourselves: “If that is the answer, just what was the question?”¹²¹

At the other end of this spectrum sits the prospect that analysts produce assessments to please the preferences and interests of policy-makers, seeking their approval by telling consumers what they believe they wish to hear.¹²²

Studies of politicization suggest that producers rarely apply direct pressure to analysts, and analysts rarely engage in blatant pandering. As a result, even the most contested political issues are unlikely to provide clear examples policy makers directly pressuring intelligence analysts and managers to bend to their will.¹²³ However, the privileged position of consumers in their relationship with producers affords them multiple opportunities to shape intelligence products through the use of their institutional prerogatives, such as calling for independent assessments, creating new analytic units, questioning analysts about specific data sources and methods, shaping analytic and

¹²¹ Gregory F. Treverton, “Intelligence Analysis: Between ‘Politicization’ and Irrelevance,” in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 92.

¹²² James J. Wirtz, “Intelligence to Please? The Order of Battle Controversy During the Vietnam War,” *Political Science Quarterly*, Vol. 106, No. 2 (Summer, 1991), pp. 239-263.

¹²³ Paul R. Pillar, “The Perils of Politicization,” in Lock K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), p. 477.

collection priorities, or even threatening producers with the loss of access to policy makers in the future or on other topics.¹²⁴

Two significant institutional reforms were made during Robert Gates' tenure as Deputy Director of Intelligence (DDI) and Director of Central Intelligence (DCI) in the 1980s and early 1990s. These reforms sought to address institutional problems between producers and consumers, and proved to be amongst the most contentious periods in the CIA's analytic history regarding the relationships between intelligence analysts, managers, and consumers. First, Gates encouraged embedding intelligence analysts in the staff of policy-making organizations; second, he insisted on putting analytic products through a coordination process of peer and managerial review prior to dissemination to consumers. As Treverton noted, sending intelligence analysts to work on policy-makers' staffs afforded them new opportunities to better understand how consumers approached problem solving and evaluated intelligence products.

One [of Gates' reforms] was sending DI [Directorate of Intelligence] analysts more frequently to serve on rotations in policy agencies—State, the NSC [National Security Council], the Pentagon, the U.S. trade representative, and elsewhere. While on such rotations, they are all-purpose staffers who happen to have intelligence expertise, just as other staffers might be lawyers or economists. They acquire a feel for the pace and rhythm of policy, as well as personal contacts within the policy world that can be acquired no other way. The effect of those rotations on intelligence officers is striking, all the more so because many DI analysts

¹²⁴ For discussions see James J. Wirtz, "Intelligence to Please? The Order of Battle Controversy During the Vietnam War," *Political Science Quarterly*, Vol. 106, No. 2 (Summer 1991), pp. 239-263; Gregory F. Treverton, *Reshaping National Intelligence for an Age of Information* (New York, NY: Cambridge University Press, 2003), p. 198; Paul R. Pillar, "Intelligence, Policy, and the War in Iraq," *Foreign Affairs*, Vol. 85, No. 2 (March/April 2006), pp. 15-27; Paul R. Pillar, "The Right Stuff," *The National Interest*, No. 91 (September/October 2007), pp. 53-59; Paul R. Pillar, *Intelligence and U.S. Foreign Policy: Iraq, 9/11, and Misguided Reform* (New York, NY: Columbia University Press, 2011); and Joshua Rovner, *Fixing the Facts: National Security and the Politics of Intelligence* (Ithaca, NY: Cornell University Press, 2011).

have spent years trying to help policy without ever serving among those who try to make it.¹²⁵

While embedding intelligence analysts in policy-making organizations gave them a new appreciation for how to be relevant to the needs of consumers, it also challenged the intelligence community's value of policy neutrality and independence. Many analysts resisted this change, and worried that working in close proximity to consumers and sustained engagements with their staff would politicize the intelligence they provided. Despite early resistance, this reform has proven to significantly strengthen the relationships between producers and consumers, and has become an important part of developing professional analysts. John McLaughlin summarized the benefits of direct, personal experience with consumers, noting it provided analysts with an improved understanding of what policy makers wanted to know, already knew, and were currently doing to address contemporary problems.

In my personal experience, however, the most effective way for analysts to understand what policymakers need is to live and work among them for a period of time. It pays enormous dividends in mutual understanding to deploy some portion of the analyst workforce on temporary rotational assignments into the policy community. These can range from assignments of several months' duration supporting an overseas embassy to a year-long stint in one of the executive branch agencies....

...While serving [in the State Department] as a special assistant to a senior officer in the Bureau of European Affairs, I had an inside look at how the CIA's work was received on a wide range of issues. I heard our work both praised and scorned and sought to understand why it sometimes elicited the latter reaction. It seldom had to do with the narrow substance of the message.

Such negative reactions more often had to do with other problems. For instance, the State Department officer had already read the raw reporting on which the analysis was based and found little new in what we

¹²⁵ Gregory F. Treverton, *Reshaping National Intelligence for an Age of Information* (New York, NY: Cambridge University Press, 2003), p. 205.

wrote. Or a particular assessment was simply too long and complicated for a harried policy officer to absorb. Or the analysis was written without a clue as to what policymakers were thinking or doing about the problem and therefore appeared naive, abstract, or uninformed. Or the analyst had pointed out all the problems surrounding an issue but paid no attention to what points of leverage or opportunities the United States might have.

There were obvious learning points in all this. To succeed, the analysis had to be timely, digestible, and informed about the policy context while stopping short of pandering to or prescribing the policy, and it needed to help policymakers in their search for leverage. Ideally, analysts serving in positions like the one I held should not be in policymaking positions but instead serve as onsite analytic resources for policymakers, with the capacity and authority to reach back into the intelligence agencies for analytic support. Analysts who have this experience gain a keen appreciation for the qualities that their work must possess in order to be taken seriously and have an impact.¹²⁶

Policy makers also benefitted from the presence of intelligence analysts on their staff, although they too were challenged by the cultural differences between the organizations. For example, intelligence analysts were quick to provide factual information and assessments of current and future events, but were regarded as unreliable on matters of policy advocacy. However, once consumers understood the institutional roles of intelligence, they benefitted greatly from having embedded intelligence producers who could identify when policy makers were relying on questionable assumptions or estimates about foreign actors that merited closer inspection, and help consumers access the larger intelligence community in order to ensure that they capitalized on the wealth of available collection and analytic resources.¹²⁷ James Steinberg noted the mutual benefits of embedded analysts, but also the institutional

¹²⁶ John McLaughlin, "Serving the National Policymaker," in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 74-75.

¹²⁷ Interview with Leon Fuerth, National Defense University, September 21, 2011.

strains that developed when policy makers evaluated their staff on matters such as political loyalty or reliability:

...[the need for] policymakers to keep intelligence representatives “in the room” when policy is debated. Although analysts rightly take a vow of silence with respect to policy prescriptions, they need to hear the underlying assumptions and beliefs that inform policy, both to correct errors of fact that may creep into policy and to provide policymakers with insights into the factors that might lead them to question or change those assumptions as events unfold. The real danger in the ongoing debate about the danger of “politicizing” intelligence is that both sides will overreact and create a “Chinese wall” that cuts off the analysts from firsthand access to policy debates. McLaughlin suggests one way to achieve this goal—namely, to embed more analysts in policymaking units, not as policymakers themselves but as part of the day-to-day activities of key agencies. For policymakers to gain the benefit of such embedded analysts, they need to appreciate and respect the fact that these analysts are different from other members of the policymaking team and thus should not be subject to the same tests of loyalty or ideological affinity that may be appropriate for “political” appointees—and even more, should not be punished or ignored for putting forth skeptical perspectives or inconvenient truths.¹²⁸

A second reform introduced by Gates involved changing the way intelligence products were produced. The introduction of a formal coordination process transformed analysis from being the highly personal, individual products of single experts into peer-reviewed assessments that represented the views of the entire organization. This change challenged the culture of independence and expertise, which analysts believed applied to their individual efforts rather than the institution itself. As Betts noted, the introduction

¹²⁸ James B. Steinberg, “The Policymaker’s Perspective: Transparency and Partnership,” in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 87-88.

of the coordination process explicitly acknowledged that, “Intelligence products are supposed to represent the best judgments of whole organizations, not single authors.”¹²⁹

The tension between analysts and managers resulted in competing narratives where analysts believed that the coordination process served as a vehicle for censoring their analysis in order to politicize intelligence assessments, while managers believed that they were ensuring greater consistency, confidence, expertise, and relevance in analytic products, thus elevating them in the eyes of consumers, and protecting the institution from consumers’ cherry picking particular analysts for particular problems. By viewing producer/consumer relations through a corporate lens, Gates argued that intelligence products are collectively produced and peer and managerial reviews provided consumers with confidence in their assessments because of its institutional imprimatur rather than individual author. Nevertheless, the charges of politicization were so fierce that Gates personally responded to them in an agency-wide address:

...unwarranted concerns about politicization can arise when analysts themselves fail to understand their role in the process. We do produce a corporate product. If the policymaker wants the opinion of a single individual, he or she can (and frequently does) consult any one of a dozen outside experts on any given issue. Your work, on the other hand, counts because it represents the well-considered view of an entire directorate and, in the case of National Estimates, the entire Intelligence Community. Analysts themselves must play a critical role in making the system work. They must do their part to help foster an open environment. Analysts must understand and practice the corporate concept. They must discard the academic mindset that says their work is their own, and they must take into account the views of others during the coordination process.¹³⁰

¹²⁹ Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York, NY: Columbia University Press, 2007), p. 81.

¹³⁰ Robert M. Gates, “Guarding Against Politicization,” *Studies in Intelligence*, Vol. 36, No 1 (1992), p. 8.

Just as embedding analysts within policy-making organizations was initially met with resistance, only to be seen as enriching the relationship between producers and consumers later, the shift toward corporate production transformed from being viewed as a subtle form of politicization into an important hedge against it. This transition was evident during the period before the 2003 Iraq War when policy makers were questioning whether a link between Saddam Hussein's regime and al-Qaeda (AQ) existed. During this time, the CIA's Office of Near East and South Asia (NESA) complained to the CIA Ombudsman for Politicization, a position created in 1992 as a result of allegations of politicization during Gates' confirmation hearings for the position of DCI, that another office, the Office of Terrorism Analysis (OTA), had undermined their position that no links between AQ and Iraq existed by producing *Iraq and al-Qaida: Interpreting a Murky Relationship*. This document explored the justification for determining a link between AQ and Iraq based on the available intelligence. NESA argued that OTA's assessment was not properly coordinated and reached unjustifiable conclusions that undermined NESA's analysis. While the Ombudsman ultimately determined that OTA's assessment was not politicized, the fact that analysts and managers argued that the coordination processes had become the mechanism by which the integrity of the production process was preserved marked a dramatic institutional change compared with the events of a decade earlier.¹³¹

¹³¹ U.S. Senate, Senate Select Committee on Intelligence, *Report on the U.S. Intelligence Community's Prewar Intelligence Assessments on Iraq* (Washington, DC: United States Senate, 7 July 2004), pp. 359-360.

More recently, the coordination process introduced by Gates has been extended as part of the post-9/11 and Iraqi WMD intelligence reforms. Pillar noted that in the production of National Intelligence Estimates (NIEs), collectors are now included in reviews of reports in order to ensure that intelligence data is not employed out of context or assigned undeserved weight with respect to the veracity of the source or method.¹³² This revised NIE production process and the role of collectors in it was explained by Fingar as a matter of ensuring that analysts have a better understanding of source validity and collection circumstances when constructing the community's flagship summaries of strategic and policy issues.

While analysts are revisiting previous judgments, collectors must rescrub the intelligence germane to the new estimate. This begins with a revetting of the reporting used to support previous judgments as well as reexamining reporting that analysts indicate they consider important to the new assessment. This, too, reflects lessons learned or relearned from the Iraq WMD estimate. Collectors—via a formal letter signed by the head of the agency or designee—must attest to the validity of the sources used in the estimate. George Tenet initiated this requirement, and its application was expanded by the ODNI [Office of the Director of National Intelligence]. The process is time consuming but critical and has become a routine part of the NIE process.¹³³

Although the coordination process has ensured that analytic products more accurately present the organization's perspective, consumers have noted that this effort can sometimes limit the timeliness of intelligence. For example, Leon Fuerth argued that as the pace of events increases and the timelines of policy makers diminish, the need for analysts to respond to questions quickly often overwhelms the coordination process.¹³⁴

¹³² Interview with Paul Pillar, Georgetown University, February 1, 2012.

¹³³ Thomas Fingar, *Reducing Intelligence: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), p. 83.

¹³⁴ Interview with Leon Fuerth, National Defense University, September 21, 2011.

He argued that once policy makers have confidence in the intelligence analyst supporting them, they trust the analyst to provide the intelligence community's perspective on pressing matters without the need for formal coordination. The continued importance of personal relationships was also affirmed by Fingar, who led the Office of the Director of National Intelligence's (ODNI's) analytic efforts to restore policy makers' confidence in the community after the Iraqi WMD estimate of 2002. Fingar noted that in spite of many methodological and procedural reforms designed to improve analytic rigor, consumers continued to evaluate intelligence products based on their trust and relationships in the individual analysts responsible for providing it to them. While recounting his experience presenting an NIE on Iran's nuclear program, Fingar noted that the highly personal act of providing intelligence resulted in the individual analyst receiving undue credit for the collective efforts of the community.

In the weeks after release of the Iran NIE, I briefed our findings to several committees and individual members.... After the first of these briefings, one of the members reached across the table, took my hand, and said, "Thank you for your courage and integrity. The American people owe you a debt of gratitude." I did not know whether he meant me personally or intended "you" to be a collective noun embracing all who had worked on the NIE but hoped it was the latter. The first time, the comment was gratifying, but it was repeated, with only slight variations, after almost every briefing and came from members of both parties and in both houses. Each repetition left me more discomforted because the statements implied that integrity was rare and heroic when I believe that it is the coin of the realm in the Intelligence Community. If we lack or are not seen to have integrity, we are little more than a useless waste of taxpayer money. Despite all the effort and all the progress that I believe we had made to improve the quality of analytic products, even those who "liked" our conclusions often attributed them to personality rather than professionalism. I did and still do find this very disappointing.¹³⁵

¹³⁵ Thomas Fingar, *Reducing Intelligence: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), p. 124.

In these cases, both Fuerth and Fingar each reaffirmed the central institutional tradeoffs that Gates' reforms sought to address: improved interpersonal relationships between producers and consumers as a means of improving the relevance and credibility of analytic products in exchange for the depersonalization of analytic production in order to improve its accuracy and authority.¹³⁶

Cultural Differences

Intelligence scholars and practitioners have often regarded producers and consumers as belonging to different cultures or “tribes,” where each group maintains different values and beliefs about the world and their place within it.¹³⁷ The culture of policy making is fundamentally activist and committed to shaping the world and events within it. As a result, they believe in the importance of their choices and seek the inputs of those whom they believe are committed to their success. Thus, consumers measure the contributions of intelligence producers by the extent to which they believe analysts support the successful determination and implementation of their policy choices.

Alternatively, the culture of analysis is based on the development and promotion of deep, substantive expertise, where analysts focus on the nuance, detail, and complexity of particular issues; strive to remain neutral on policy preferences; and view their contributions to consumers in protective terms—warning policy makers when their strategies and policies may rest on unwarranted assumptions or be overtaken by

¹³⁶ In this context, accuracy does not mean predictive accuracy. Instead it refers to the extent to which an analytic product accurately reflects the available intelligence and the perspectives of those in the intelligence community.

¹³⁷ Mark M. Lowenthal, “Tribal Tongues: Intelligence Consumers, Intelligence Producers,” *The Washington Quarterly*, Vol. 15, No. 1 (1992), pp. 157-168.

events.¹³⁸ As Lowenthal noted, producers and consumers clash because they possess divergent beliefs about what it means to support policy and policy makers.

Most policymakers (i.e., consumers) work on the assumption of basic support throughout the government for their various policy initiatives, including support by the intelligence community. The first problem lies in the very word *support*. For policymakers, this means a shared and active interest and, if necessary, advocacy. This runs counter, however, to the intelligence community's long-standing position not to advocate any policy. Rather, the intelligence community tends to see itself, correctly or not, as a value-free service agency, although at its upper levels the line begins to blur.

Second, the intelligence community, like all other parts of the permanent government bureaucracy, has a "we/they" view of its political masters. The intelligence community is part of the permanent government; those making policy are politically driven transients, even when nominated from within the professional ranks of agencies. Indeed, with the exception of the uniformed military, nowhere else in the entire foreign policy and defense apparatus can there be found as many career officials at such senior levels as in the intelligence community.¹³⁹

These cultural differences between producers and consumers are often framed as optimists vs. skeptics or pessimists. Policy makers are referred to as optimists because of their belief that their actions and decisions can change the world. As Lowenthal noted, policy-makers "...approach problems with the belief that they can solve them. After all, this is the reason they have gone into government."¹⁴⁰ Because they believe their choices matter, bureaucratic and organizational politics have real consequences, incentivizing the competition for control over agendas and resources. Thus, intelligence support is often viewed through a bureaucratic lens.

¹³⁸ James B. Bruce and Roger Z. George, "Intelligence Analysis—The Emergence of a Discipline," in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 9.

¹³⁹ Mark M. Lowenthal, "Tribal Tongues: Intelligence Consumers, Intelligence Producers," *The Washington Quarterly*, Vol. 15, No. 1 (1992), pp. 157-158.

¹⁴⁰ Mark M. Lowenthal, *Intelligence: From Secrets to Policy* (Washington, DC: Congressional Quarterly Press, 2009), p. 186.

The different cultures of policy and intelligence hold the potential to inject a great deal of misunderstanding and tension into the relationship. The culture of the policy world is marked by elements of realism but is essentially—and necessarily—a culture of optimism. Policymaking is a contentious business marked by lots of competing ideas and frequently by heavy intellectual combat. A lot of bureaucratic blood is often on the ground once a particular course of action wins out.¹⁴¹

As a result of domestic politics, policy makers rarely view intelligence products as neutral, instead interpreting them as weapons to be employed or neutralized within the context of domestic politics. In this sense, intelligence is necessarily political because the policy makers it supports are engaged in a political process. Thus, the lens by which intelligence products are evaluated is not substantive, but rhetorical. The irony, then, is that consumers view intelligence that does not support their position or goal as politicized. As Paul Wolfowitz noted, policy makers looking to use intelligence analysis often view the producer's culture of neutrality as a form of politicization that places consumers into an "accept or reject" position when viewing intelligence products.

Perhaps this is revealing a certain arrogance on my part, but I frequently think I am as capable of coming up with an informed opinion about a matter as any number of the people within the Intelligence Community who feel that they have been uniquely anointed with this responsibility. Too often that attitude is a doge that allows them to conceal ignorance of facts, policy bias, or any number of things that may lie behind the personal opinions that are presented as sanctified intelligence judgments. No one is allowed to question those judgments, especially not policymakers, or, as the argument goes, they will pollute the intelligence process with policy judgments.

I think that this attitude on the part of the Intelligence Community causes a lot of problems. I think that it actually encourages the manipulation of intelligence judgments for political policy purposes. If you can get the authority of the Intelligence Community on your side, you

¹⁴¹ John McLaughlin, "Serving the National Policymaker," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 71-72.

can appeal to authority without having to bother appealing to the evidence. More important, it places great importance on a product that reports the judgments of analysts, which, absent the evidence on which those judgments rest, have limited value to policymakers. It tends to produce turgid National Intelligence Estimates (NIEs), marked by summary judgments in the front, full of carefully balanced sentences (“on the one hand,” and “on the other hand”), offering no new facts or reasoning to which any sophisticated reader of the weekly reader would not already have access. On a busy day, they are not even looked at. They may be glanced at because you have to know what someone may say to you at a meeting, citing the judgment of the NIE as an authority. Yet you do not read them expecting to learn very much. An estimate may be a useful weapon in a debate or it may be someone else’s weapon against which you have to be prepared to defend yourself, but you rarely read them expecting to learn something new.¹⁴²

From the policy makers’ perspective, intelligence is accepted or rejected based on its ability to support the consumer’s ability to advocate for preferred policies. Moreover, as Wolfowitz noted, policy makers prefer to examine and interpret intelligence information for themselves—setting aside intelligence analysts’ substantive judgments and evaluations of sources in favor of their own interpretation.

The view of intelligence from the consumer’s perspective that Wolfowitz offered is particularly significant. While the passage above was written in 1995, it foreshadowed the precise context in which he evaluated intelligence reports regarding Iraq’s WMD capabilities and ties to AQ while serving as the Deputy Secretary of Defense in the first George W. Bush administration—a period that included several controversial analytic and bureaucratic maneuvers in order to marshal support for invading Iraq based on highly disputed intelligence sources and interpretations. Indeed, the essentially bureaucratic character of intelligence assessments was noted in an interview with *Vanity Fair* in May

¹⁴² Paul Wolfowitz, “Comments: Paul Wolfowitz,” in Roy Godson, Ernest R. May and Gary Schmitt, eds., *U.S. Intelligence at the Crossroads: Agendas for Reform* (Washington, DC: Brassey’s, 1995), p. 76.

2003, when Wolfowitz stated that although there were several reasons to remove Saddam Hussein from power, the bureaucracy could only agree on WMD as a *causus belli*, thus establishing the context in which intelligence products were interpreted and valued.

[Vanity Fair] Q: Was that one of the arguments that was raised early on by you and others that Iraq actually does connect, not to connect the dots too much, but the relationship between Saudi Arabia, our troops being there, and bin Laden's rage about that, which he's built on so many years, also connects the World Trade Center attacks, that there's a logic of motive or something like that? Or does that read too much into—

Wolfowitz: No, I think it happens to be correct. The truth is that for reasons that have a lot to do with the U.S. government bureaucracy we settled on the one issue that everyone could agree on which was weapons of mass destruction as the core reason....¹⁴³

The views expressed by Wolfowitz and other policy makers suggest that the optimism of the policy culture is not a casual one. Policy makers do not believe that events will play out to their benefit absent concerted, focused efforts on their part to intervene in the international system. As a result, their optimism is really an expression of agency and their belief that they reside in a competitive, strategic system that constantly begs action. Because their choices matter, the ways in which they think about problems, define success and failure, and evaluate options become central concerns for mobilizing and sustaining support for specific policies. Thus, the policy world is primarily and necessarily a world of ideas and ideology regarding future possibilities rather than facts and data about the past and present.

In contrast to policy makers, intelligence analysts are regarded as skeptics or pessimists who overwhelmingly focus on uncovering and describing the ways in which

¹⁴³ Paul Wolfowitz, *Deputy Secretary Wolfowitz Interview with Sam Tannenhaus, Vanity Fair*, May 9, 2003, <http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2594> (accessed on May 30, 2012).

policy makers' plans may fail, be countered by the strategic moves of others, or be overtaken by events.

In contrast to the fundamentally optimistic thrust of the policy culture, the culture of the intelligence world is marked by skepticism. The requirement to warn of dangers—and the heavy criticism when warning fails—encourages a darker view than is ever instinctively the case in the policy world. A former director of central intelligence used to define an intelligence analyst as someone who “smells flowers...and then looks for the coffin.” In short, analysts are trained, and indeed are required, to look for trouble—regrettably often at the expense of opportunity—and are thought to have failed when they do not detect it.¹⁴⁴

The constant concern over policy failures and threats places intelligence producers in a competitive posture vis-à-vis consumers, who often resist analysis that suggests the results of their hard-fought bureaucratic battles might fail. While intelligence analysts believe that their skepticism protects policy makers from overconfidence and wishful thinking in the face of immutable uncertainties, consumers often perceive their warnings as a form of criticism and a direct challenge to their agendas.

Given their perspective, intelligence analysts can easily fall into thinking that part of their job is to protect overeager policy officials from their own enthusiasms. Of course, those policy officials see just the reverse: They see intelligence as perennial naysayers, eager to stick a finger in the eye of policy.¹⁴⁵

Numerous studies of producer/consumer relations have noted that policy makers generally appreciate the delivery of factual information and updates on current events, basic and current intelligence, but often dismiss the judgments of analysts whenever they speculate about uncertain futures or information beyond their empirical reach. More than

¹⁴⁴ John McLaughlin, “Serving the National Policymaker,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 73.

¹⁴⁵ Gregory F. Treverton, *Reshaping Intelligence for an Age of Information* (New York, NY: Cambridge University Press, 2003), p. 181.

five decades ago Roger Hilsman surveyed many intelligence consumers and concluded that the very act of offering warnings about potential threats served as an implicit criticism of current policy, thus straining relations between producers and consumers.

One kind of warning, for example, could be given when a threat is not foreseen because no one even suspects that there may be a problem. In the past such a complete failure to recognize a problem has usually occurred when the necessary conceptual tools simply did not exist. The Great Plague is a case in point. Without knowledge of the cause and means of transmittal of disease, the filth and increase in rats which made the plague possible had no significance. Obviously, in such circumstances it would be unreasonable to expect much from intelligence. Another kind of warning could be given when the operators have recognized the existence of a problem, but have decided after analysis that it will work itself out without harmful effects. If intelligence warns that there is a threat after all, it is in effect criticizing the analysis of the people who decided otherwise, and saying that the present policy of taking no action to meet the threat is wrong. A third type of warning could be given when a threat has already been recognized, and a decision to employ a certain means already made. In these circumstances, a warning that the threat will materialize after all implies that the chosen means is ineffective, or that it will bring new threats—again, that the policy is wrong. And notice that an intelligence unit could not justify a statement that the present course was not the proper one unless it had made a detailed analysis as well as a preliminary one. The “warning” role, in sum, will usually become a “warning-critic” one.¹⁴⁶

Hilsman’s findings on the problems of warning were echoed by Kent, who noted that policy makers reported that estimates whose conclusions consumers agreed with were regarded as a waste of their precious time, while those they disagreed with were either dismissed or attacked.¹⁴⁷ Six decades later, Pillar noted the same problems existed and, when compounded by institutional authority, could lead to the politicization of intelligence production.

¹⁴⁶ Roger Hilsman, Jr., “Intelligence and Policy-Making in Foreign Affairs,” *World Politics*, Vol. 5, No. 1 (October 1952), pp. 34-35.

¹⁴⁷ Sherman Kent, “Estimates and Influence,” *Studies in Intelligence*, Vol. 12, No. 3 (1968), pp. 18-19.

In the...real world of politics and policy making, decision-makers more commonly arrive early at their own conclusions and devote most of their attention to the sometimes difficult task of mobilizing support for the policies they have selected. Anything that makes that task even more difficult is likely to annoy or anger them. Knowing this is a powerful influence on anyone, including intelligence officers, who work for the policymakers.

Displeasing the policymaker, through intelligence products that make his political task harder rather than easier, can spoil an intelligence officer's day in numerous ways. The cost can be as simple as a critical or biting remark, which, if coming from a powerful person, can be a major blow to a relatively powerless one. The cost may take the more pointed form of accusations that the intelligence officers involved are not team players and are not supporting policymakers as they are supposed to. The costs may be especially acute for the most senior intelligence officers, who must deal directly with policymakers, regularly and face-to-face. They are likely to feel the most pain from any suggestion that they are not team players, because to do their job they to some extent are co-opted onto the policy team. The specific sanctions may include exclusion from the policy-making circle, making them even more ineffective and irrelevant, or loss of their positions altogether. Whatever is the politicizing effect on senior intelligence officers, a ripple effect is felt down through the organizations that they lead.¹⁴⁸

The perceived pessimism of intelligence analysts often exacerbates the existing institutional tensions between producers and consumers. By constantly warning of impending failures and treats, intelligence analysts are seen as the bearers of bad news who are not team players at best, and actively working against the interests of their policy superiors at worst. President Lynden Johnson colorfully described the intelligence community as a culture committed to undoing the hard work of policy makers rather than assisting them.

Let me tell you about these intelligence guys. When I was growing up in Texas we had a cow named Bessie. I'd go out early and milk her.... One day I'd worked hard and gotten a full pail of milk, but I wasn't paying

¹⁴⁸ Paul R. Pillar, "The Perils of Politicization," in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), pp. 477-478.

attention, and old Bessie swung her shit-smeared tail through that bucket of milk. Now, you know, that's what these intelligence guys do. You work hard and get a good program or policy going, and they swing a shit-smeared tail through it.¹⁴⁹

Consumers may also dismiss intelligence warnings on grounds other than optimism. Often, policy makers believe that they do not have the luxury of waiting for more information or doing nothing as a potential crisis unfolds and must often choose between many sub-optimal or bad alternatives. As a result, rather than see policy and strategy as the search for an optimal solution to a problem, they understand that they may need to engage a problem on unfavorable terms and learn, adapt, and evolve solutions.¹⁵⁰ In such circumstances, warnings of policy failures may be perceived as unhelpful since what is needed are greater insights into the relative merits of available options, rather than simply accounting for the downsides of each.¹⁵¹

...there is a perception by policymakers that the analytic community views its role as one of cautioner (or worse, naysayer) rather than a support to policy. McLaughlin refers to this as the policymakers' culture of optimism versus the analysts' culture of skepticism. Another way of thinking about this is that policymakers rarely have the luxury of throwing up their hands and saying "too hard" or deferring decisions until the intelligence becomes clearer; often they must act even if the choices are muddy and the consequences are unpredictable.¹⁵²

Efforts to address the cultural differences between producers and consumers have started to redefine the boundaries between policy making and intelligence support,

¹⁴⁹ Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York, NY: Columbia University Press, 2007), p. 24.

¹⁵⁰ Thomasingar, *Reducing Uncertainty: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), pp. 112 and 15-16.

¹⁵¹ Michael Lewis, "Obama's Way," *Vanity Fair*, October 2012, <http://www.vanityfair.com/politics/2012/10/michael-lewis-profile-barack-obama> (accessed on September 17, 2012).

¹⁵² James B. Steinberg, "The Policymaker's Perspective: Transparency and Partnership," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 83-84.

specifically with respect to the identification of opportunities for policy makers to act in order to influence events within the international system. As James Steinberg noted, the identification of opportunities simultaneously recasts intelligence analysts as the bearers of bad news while challenging the analytic community's self-limiting definition of neutrality.

Policymakers look to the intelligence community to uncover the facts that will help them achieve their goals. Contrary to the views of some critics, most policymakers do not resist bad news if it is reliable and timely, because they know they cannot succeed by sticking their heads in the sand and pretending that adverse developments will go away if they simply ignore or dismiss them. But often policymakers feel that the intelligence community views its mission as solely being the bearer of bad news or "warning"—that is, telling the policy community about all the obstacles to achieving their objectives, rather than identifying opportunities and how to make the best of the situation to achieve them. Yet for many analysts, such a role is tantamount to "supporting" the policy and thus violating the most sacred canon of analytic objectivity and policy neutrality.¹⁵³

Efforts to improve relations between producers and consumers have encouraged intelligence analysts to play a stronger hand in the formulation of policy and strategy by identifying and informing policy makers about opportunities to shape the international system.¹⁵⁴ The introduction of opportunity analysis into the analytic community marked another evolutionary step in producer/consumer relations by urging analysts to actively assist in the achievement of policy makers' objectives. The significance of the changes created by opportunity analysis were noted by John Hedley, and summarized as a

¹⁵³ James B. Steinberg, "The Policymaker's Perspective: Transparency and Partnership," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 84.

¹⁵⁴ John H. Hedley, "The Evolution of Intelligence Analysis," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, Innovations* (Washington, DC: Georgetown University Press, 2008), p. 30.

departure from “old-school” analytic production that was far more removed from policy and therefore less actionable and useful to consumers.

The desire on the part of the users of intelligence is for analysis that is opportunity-oriented, or actionable—in other words, intelligence they can apply and actually use. Analysis has become an integral part of planning and implementing policy, and of intelligence operations. This is a far cry from what might be termed the traditional, or “old school,” conception of analysis which held that to earn and maintain credibility, analysts must be more than policy neutral, they must literally keep their distance from those who were making the policy decisions. Traditional thinking also held that analysis should—again, to earn and maintain credibility—be done independently of those who collect it.¹⁵⁵

Fingar expanded on Hedley’s points by further contextualizing the biases within the intelligence community that assigned highest priority to warning and threats over opportunities for shaping.

A substantial portion of what we spend to reduce uncertainty—more than \$50 billion a year—goes to the Intelligence Community. The need for this amount of money is justified through a process that emphasizes threats. For example, the classified and unclassified versions of the Annual Threat Assessment submitted to the Congress by the director of national intelligence devote far more attention to problems and perils than to opportunities for positive change. This emphasis is understandable, but it is also unfortunate because it obscures one of the most important functions of the Intelligence Community and causes both analysts and agencies to devote too little attention to “good news” and potential opportunities to move developments in a more favorable direction.¹⁵⁶

While many in the analytic community believed that opportunity analysis crossed the line from neutrality to advocacy, it was something that consumers welcomed and expected from intelligence producers.

¹⁵⁵ John Hollister Hedley, “Analysis for Strategic Intelligence,” in Loch K. Johnson, ed., *Handbook of Intelligence Studies* (New York, NY: Routledge, 2009), p. 218.

¹⁵⁶ Thomas Fingar, *Reducing Uncertainty: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), pp. 50-51.

Policy makers understand that I&W [Indicators and Warnings] leaves them in the position of reacting to intelligence. But policy makers also want to be actors, to achieve goals and not just prevent bad things from happening. As more than one senior policy maker has said, “I want intelligence that helps me advance my agenda, that makes me the actor, not the reactor.” This is often referred to as **opportunity analysis**.¹⁵⁷

Opportunity analysis requires strengthening the relationship between producers and consumers, and like the reforms implemented under Gates’ leadership tenure, require the development of strong personal relations and trust between analysts and policy makers.¹⁵⁸ Opportunity analysis has become a mechanism for building personal trust between producers and consumers by changing the cultural dynamics of their interactions. Providing notification of unexpected opportunities to advance their agendas and achieve their goals has been a welcomed development for consumers, and has encouraged analysts to play a more constructive role in the formulation of strategy and policy without offering prescriptions.¹⁵⁹

Opportunity analysis (sometimes referred to as action analysis) directly supports implementation of U.S. security policies by assessing the factors that could help policy planners and other decision makers to seize opportunities presented to them. While not endorsing any policy options, opportunity analysis assesses the costs and benefits of different policy actions that policymakers might consider.¹⁶⁰

¹⁵⁷ Mark M. Lowenthal, *Intelligence: From Secrets to Policy* (Washington, DC: Congressional Quarterly Press, 2009), p. 134.

¹⁵⁸ Thomasingar, *Understanding Uncertainty; Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), pp. 50-52.

¹⁵⁹ Jack Davis, “A Policymaker’s Perspective On Intelligence Analysis,” *Studies in Intelligence*, Vol. 38, No. 5 (1995), p. 11; and Roger Z. George, “The Art of Strategy and Intelligence,” in Roger Z. George and James B. Bruce, *Analyzing Intelligence: Origins, Obstacles, Innovations* (Washington, DC: Georgetown University Press, 2008), p. 109.

¹⁶⁰ Roger Z. George and James B. Bruce, “Glossary of Analysis Terms,” in Roger Z. George and James B. Bruce, *Analyzing Intelligence: Origins, Obstacles, Innovations* (Washington, DC: Georgetown University Press, 2008), p. 316.

In the context of producer/consumer relations, successful opportunity analysis constitutes the pinnacle of intelligence analysts and policy makers working in harmony. The successful production and acceptance of opportunity analysis requires producers to be taken into the confidence of consumers: Analysts must be privy to policy makers' goals, plans, and worldviews in sufficient detail as to look over the horizon on their behalf. Moreover, given the intelligence community's coordination process, successful opportunity analysis also means that analysts' peers believe that assessments have remained analytic and not crossed over into advocacy. The difficulty of reaching this level of institutional and cultural accord was noted by Lowenthal, who argued:

Opportunity analysis is a sophisticated but difficult type of analysis to produce. First, it requires that the intelligence managers or analysts have a good sense of the goals that the policy maker seeks to achieve. Successful opportunity analysis may require some degree of specific and detailed knowledge of these goals. For example, knowing that a goal is arms control may not suggest many useful avenues of opportunity analysis beyond broad generalities. Knowing that the goals include certain types of weapons or restrictions would be more helpful. Thus, again, emphasis is placed on the importance of the intelligence analysts knowing the intended directions of policy. Second, opportunity analysis often seems more difficult or riskier as it requires positing how foreign leaders or nations will react to policy initiatives. Positing a foreign action and then describing either the consequences or possible reactions often seems easier than the reverse process. After all, an analyst often feels more comfortable understanding how a nation or its policy makers are likely to react even if the analyst is an expert in the politics of another country. Finally, opportunity analysis brings the intelligence community close to the line separating intelligence from policy. Writing good opportunity analysis without appearing to be prescriptive can be difficult even if that is not the intended message or goal.¹⁶¹

¹⁶¹ Mark M. Lowenthal, *Intelligence: From Secrets to Policy* (Washington, DC: Congressional Quarterly Press, 2009), p. 134.

Epistemology of Intelligence Production

The epistemological foundation of intelligence analysis leaves it vulnerable to criticism, and always provides consumers with an opportunity to dismiss producer's analytic judgments. The reason for this is because, by definition, any statement about the future state of the world is ultimately speculative and lies beyond the limitations of empirical experience and observation. As a result, policy makers can always resort to their own expertise, analysis, trusted advisors, or wishful thinking due to the fundamental limitations of intelligence products. Therefore, intelligence analysts have always recognized that their products are only one of many sources of information that consumers have access to, and must accept the fact that policy makers may prefer the perspectives from other sources for a variety of reasons. As Bruce and George noted, "Intelligence officials cannot control which sources of information policymakers will use or how they will use them—that is the sole prerogative of policymakers."¹⁶²

Because policy makers have access to sources of information other than intelligence products, analysts must seek to distinguish their products in two ways. First, analysts have access to classified information collected via multiple intelligence collection disciplines [INTs] that give their assessments special appeal based on the uniqueness of their sources and methods, and the exclusivity of their distribution. However, the prestige that sources and methods provide is often short-lived once policy makers recognize the limits of collection and learn not to confuse secrecy and scarcity of intelligence sources with their accuracy or insight.

¹⁶² James B. Bruce and Roger Z. George, "Intelligence Analysis—The Emergence of a Discipline," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 6.

According to a former senior intelligence officer, new administrations or governments tend to start off being very impressed with what their intelligence agencies can do for them, even though many of these policymakers have served in government before, although likely in lower positions. However, over time, the policymakers become jaded and will ask to see “the good stuff,” as if the intelligence officers have been holding out on them. When they are told that they have been seeing “the good stuff,” the policymakers are disappointed.¹⁶³

A second way that analysts seek to differentiate their products from other sources is based on their individual and collective expertise. The intelligence community analytic staff resembles a research university, where analysts represent a wide variety of scholarly and professional disciplines, and often have more tenure working on particular problems than their counterparts in other government departments.¹⁶⁴ As a result, the intelligence community can provide unique insights based on the combination of its sources and methods, awareness of policy makers’ needs and interests, and community-wide expertise. By offering tailored analytic products, customized to the needs and interests of consumers, intelligence analysts differentiate themselves based on the quality of their insights rather than their information sources.¹⁶⁵

The tension between the value of intelligence being determined by the evidence analysts provide to consumers derived from many INTs, or the inferences made by analysts themselves, provides a glimpse into the structure of intelligence production and its underlying epistemology. Kent characterized the production of intelligence as akin to

¹⁶³ John McLaughlin quoted in Mark M. Lowenthal, “The Policymaker-Intelligence Relationship” in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), p. 440.

¹⁶⁴ Roger Z. George, “Central Intelligence Agency: The President’s Own,” in Roger Z. George and Harvey Rishikof, eds., *The National Security Enterprise: Navigating the Labyrinth* (Washington, DC: Georgetown University Press, 2011), pp. 159-160.

¹⁶⁵ Joshua Kerbel and Anthony Olcott, “Synthesizing with Clients, Not Analyzing for Customers,” *Studies in Intelligence*, Vol. 54, No. 4 (December 2010), pp. 11-27.

constructing a pyramid upon which assessments rested. In this metaphor, three types of intelligence products combined in order to form the structure: basic intelligence, current or reportorial intelligence, and speculative or evaluative intelligence, depicted in Figure 3-1 below.

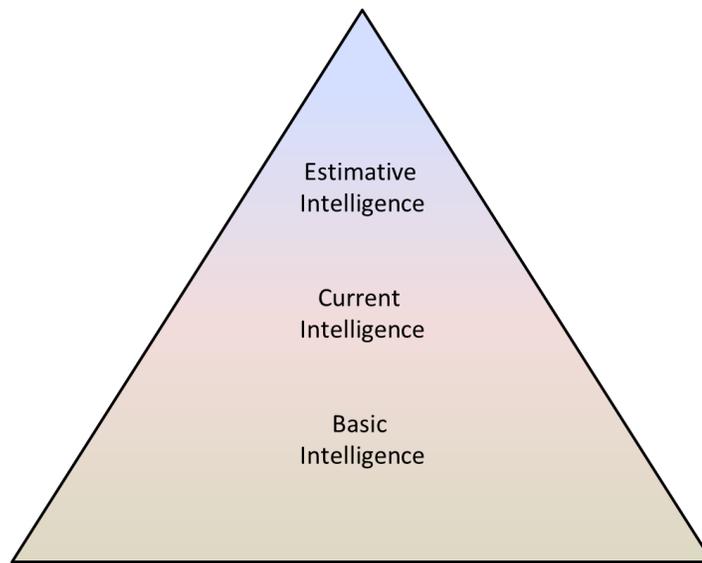


Figure 3-1: Sherman Kent's Intelligence Pyramid. Kent's intelligence pyramid shows the structure or architecture of intelligence assessments. Assessments may be characterized as estimates regarding potential future developments, or the value of missing information, that are constructed from observations of current activities, i.e. current intelligence, and verifiable, collected facts, i.e. basic intelligence. Image adapted from Sherman Kent, "Estimates and Influence," *Studies in Intelligence*, Vol. 12, No. 3 (1968), pp. 14-17.

Basic intelligence represents the collection of empirical facts: population, geography, infrastructure, climate, natural resources, order of battle, and so on.¹⁶⁶ This kind of intelligence is encyclopedic in nature. Its content is either permanent, or does not

¹⁶⁶ Sherman Kent, *Strategic Intelligence for American World Policy* (Princeton, NJ: Princeton University Press, 1949), p. 11-29.

change quickly over time, and is composed of verifiable facts. Basic intelligence forms the base of the pyramid upon which intelligence assessments or estimates rest.¹⁶⁷

Current intelligence provides descriptions of ongoing events and developments. Current intelligence concentrates on developing an understanding of the intelligence target, its current behavior, and state in a fashion akin to journalism. Current intelligence is particularly important because it serves as the primary mechanism with which intelligence analysts detect and track changes in the international system and the behavior of intelligence targets. When characterizing current intelligence, Kent noted:

As the reporting element carries out its task it constantly adds freshness to the content of the basic descriptive element. It does more than this, for in keeping otherwise static knowledge up-to-date it maintains a bridge between the descriptive and what I have called the speculative–evaluative elements—a bridge between the past and future.¹⁶⁸

Importantly, current intelligence is descriptive and reportorial rather than analytic. It links the factual base of the intelligence pyramid to its speculative apex via observation and description. Kent emphasized that as the basis of the information supporting intelligence analysis shifted from basic to current (and even more so estimative), consumers could match the expertise of producers. Whereas intelligence analysts were likely better informed than policy makers on the basic facts and history of specific problems, as considerations moved into current events and speculation about potential futures, the interested policy makers could, and often would, be as informed as their intelligence counterparts and therefore increasingly likely to perform their own analysis.

¹⁶⁷ Sherman Kent, “Estimates and Influence,” *Studies in Intelligence*, Vol. 12, No. 3 (1968), pp. 14-15.

¹⁶⁸ Sherman Kent, *Strategic Intelligence for American World Policy* (Princeton, NJ: Princeton University Press, 1949), p. 38.

The factual stuff of the base of the pyramid is likely to be largely the fruit of our own intelligence-gathering efforts and so constitute a body of material about which we are better informed than our consumers. But we enjoy no such primacy with respect to the matter above. In fact, the talent to deduce rigorously is one which we share with any other educated and intellectually disciplined human. Furthermore, the advantage we enjoy with respect to base material can be and usually is dissipated by our habit of making it available to quite an array of non-intelligence types. The point is that the studious consumer can approach our mastery at the base and match us higher up. He can be his own estimator whenever he wishes to invest the time.¹⁶⁹

The third category of intelligence, estimative or speculative, is the most challenging to employ or characterize, and is also the most problematic with respect to producer/consumer relations. Because estimative intelligence is forward looking and analytical, it rests upon the placement of facts, observations, and guesses into a theoretical framework for projecting what a target may do in the future, what options it has, and what its reactions to others' policies or actions might be. Estimative intelligence may be regarded as the search for the strategic stature, vulnerabilities, and potential of other actors.¹⁷⁰ As Kent stated, "estimating is what you do when you do not know."¹⁷¹

Because estimative intelligence abstracts from known facts and current reporting to look ahead in time, it is concerned with the limitations of what is empirically knowable. Thus, estimates about the future are more dependent on theory than facts and empirical observations.¹⁷² Therefore, the peak of the intelligence pyramid is inferential,

¹⁶⁹ Sherman Kent, "Estimates and Influence," *Studies in Intelligence*, Vol. 12, No. 3 (1968), p. 15.

¹⁷⁰ Sherman Kent, *Strategic Intelligence for American World Policy* (Princeton, NJ: Princeton University Press, 1949), p. 40.

¹⁷¹ Sherman Kent, "Estimates and Influence," *Studies in Intelligence*, Vol. 12, No. 3 (1968), p. 13.

¹⁷² Sherman Kent, *Strategic Intelligence for American World Policy* (Princeton, NJ: Princeton University Press, 1949), pp. 45-46.

derived from the patterns observed in the data and descriptions of its base, as well as other auxiliary assumptions that may be needed to fill in gaps in available information. The epistemological foundation of intelligence estimates leaves them permanently vulnerable to those who are discomforted by their conclusions. Thus, for the institutional and cultural reasons noted earlier, policy makers can always deconstruct the pyramid and construct new ones in their place based on a different theory, alternative facts, different observations, etc.¹⁷³ Therefore a paradox emerges where intelligence estimates that offer the most forward-looking assessments of potential futures are the most capable of aiding in the development and evaluation of policy options, but are also the most likely to be rejected by consumers because they necessarily consist of a series of conjectures. As George noted, “Policymakers always wanted the facts, but seldom CIA’s opinions.”¹⁷⁴

The epistemological circumstances of intelligence products ultimately provide boundary conditions against which the accuracy of assessments cannot be determined without the passage of time. Indeed, Carl von Clausewitz’s classic strategic theory expressed in *On War* provided a cautious, skeptical warning regarding the reliability of intelligence because of the epistemological circumstances surrounding its production and analysis. Like Kent, his ideas were expressed in an architectural metaphor, used to challenge the notion that decision makers could know *a priori* the quality of the intelligence they possessed at the moment of a decision.

By “intelligence” we mean every sort of information about the enemy and his country—the basis, in short, of our own plans and operations. If we

¹⁷³ Sherman Kent, “Estimates and Influence,” *Studies in Intelligence*, Vol. 12, No. 3 (1968), pp. 17-18.

¹⁷⁴ Roger Z. George, “Central Intelligence Agency: The President’s Own,” in Roger Z. George and Harvey Rishikof, eds., *The National Security Enterprise: Navigating the Labyrinth* (Washington, DC: Georgetown University Press, 2011), p. 163.

consider the actual basis of this information, how unreliable and transient it is, we soon realize that war is a flimsy structure that can easily collapse and bury us in its ruins. The textbooks agree, of course, that we should only believe reliable intelligence, and should never cease to be suspicious, but what is the use of such feeble maxims? They belong to that wisdom which for want of anything better scribblers of systems and compendia resort to when they run out of ideas....

This difficulty of *accurate recognition* constitutes one of the most serious sources of friction in war, by making things appear entirely different from what one had expected.¹⁷⁵

Challenges in the relationship between intelligence producers and consumers significantly affect intelligence analysis. Intelligence scholars and practitioners have noted that while many of the intellectual activities analysts perform resemble those of social scientists and journalists, the context in which they work is distinctive.¹⁷⁶ Many efforts have been made to improve the relationships between producers and consumers, most notably in the areas of institutional arrangements and types of analytic support offered by the intelligence community. However, these institutional efforts cannot ensure that analysis is free from bias, errors, or transparent to consumers in order improve its receptivity.

Structured Analytic Techniques and Analytic Tradecraft

The remainder of this chapter is devoted to SATs and their role in analytic tradecraft. SATs constitute the methodological counterpart to the institutional reforms introduced by Gates and the increasing emphasis on opportunity analysis. SATs developed as a result of several intelligence failures in which problems in

¹⁷⁵ Carl von Clausewitz, *In War* (Princeton, NJ: Princeton University Press, 1987), Michael Howard and Peter Paret, trans., p. 117.

¹⁷⁶ Jennifer E. Sims, "Intelligence Analysis," in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence*, (New York, NY: Oxford University Press, 2010), pp. 389-390; and interview with Paul Pillar, Georgetown University, February 1, 2012.

producer/consumers relations or failures of collection were incapable of explaining the character of faulty assessments. By introducing the notion of mindsets, the cognitive processes and capabilities of the human brain became a causal factor in intelligence failure and the target of investments in tradecraft and methodology. SATs are the product of these efforts, and serve as a midpoint on the path from traditional, intuitive analysis to a model-centric tradecraft.

The Problem of Mindsets

The history and study of intelligence failures provide the context for understanding the mindsets and their role in the intelligence community. Mindsets became a distinct topic in the study of intelligence failures that resulted from earlier research on producer/consumer relations and the interface between collection and analysis.¹⁷⁷ The institutional and cultural difficulties encountered in producer/consumer relations suggested that intelligence may fail to influence policy makers because analysts and consumers do not develop the necessary trust and personal relationships that ensure products are regarded as authoritative, informative, and relevant to the decision-making needs of policy makers.¹⁷⁸ Likewise, systemic perspectives on the international system suggested that assessments may prove faulty because of epistemological limitations

¹⁷⁷ Roberta Wohlstetter, *Pearl Harbor: Warning and Decision* (Stanford, CA: Stanford University Press, 1962); Robert Jervis, "Improving the Intelligence Process: Informal Norms and Incentives," in Alfred C. Maurer, Marion D. Tunstall, and James M. Keagle, eds., *Intelligence: Policy and Process* (Boulder, CO: Westview Press, 1985), pp. 113-124; and James B. Bruce, "The Missing Link: The Analyst-Collector Relationship," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 191-210.

¹⁷⁸ Jack Davis, *A Compendium of Tradecraft Notes* (Washington, DC: Central Intelligence Agency, February 1997), pp. 1-7; and Thomas Fingar, "Office of the Director of National Intelligence: Promising Start Despite Ambiguity, Ambivalence, and Animosity," in Roger Z. George and Harvey Rishikof, eds., *The National Security Enterprise: Navigating the Labyrinth* (Washington, DC: Georgetown University Press, 2011), p. 145.

regarding what can be predicted, and that data collection alone cannot provide an adequate basis for understanding or predicting the strategic behavior of intelligence targets.¹⁷⁹ Thus, mindsets introduced analysts' cognitive processes and psychological biases as explanatory factors in intelligence failures.¹⁸⁰

Failures to correctly anticipate world events, both large and small, have consistently demonstrated how analysts' assumptions and experiences have shaped all aspects of the intelligence process—from determining what information collectors should pursue, to how data is assessed, to what findings are deemed important enough to provide to consumers. Postmortems of these failures have often reached two conclusions: A greater devotion to strategic intelligence is necessary, and analysts must improve their use of tradecraft in order to simultaneously stimulate their imagination and increase the rigor of their assessments, confronting their own worldviews or analytic paradigms, i.e. mindsets.

Two fundamental problem areas stand out—one relatively new, or at least relatively newly revived, and one timeless. The first is a decline in the U.S. intelligence community's commitment to strategic analysis and the crafting of strategic intelligence products. This has been singled out not only in the reviews of intelligence performance on the terrorist threat leading up to September 11 but also in studies of intelligence performance on other contentious issues in the past decade.

The second problem is the enduring trap described variously as “cognitive bias” or “cognitive disassociation” or, in more derogatory terms, “mindset,” “preconception,” and “groupthink.” Taking on this issue entails treading on the sensitive terrain of inherent human

¹⁷⁹ Mark M. Lowenthal, “Intelligence in Transition: Analysis after September 11 and Iraq,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 229; and Mark M. Lowenthal, “Towards a Reasonable Standard for Analysis: How Right, How Often on Which Issues?” *Intelligence and National Security*, Vol. 23, No. 2 (June 2008), p. 313.

¹⁸⁰ Roger Z. George, “Reflections on CIA Analysis: Is it Finished?” *Intelligence and National Security*, Vol. 26, No. 1 (February 2011), pp. 72-81.

weaknesses. A common response to critiques of “cognitive bias” or “mindset” is that they are “unhelpful.” The science of “cognition” is too often regarded as academic theory. And while intelligence managers periodically attempt to make analysts aware of cognitive pitfalls, these efforts have been most commonly relegated to classrooms and have rarely made it into line practices of intelligence analysis or the evaluation of finished products.

But intelligence is a profession of cognition. This is what the job is all about—how we absorb and mentally process information coming to us, and how the receivers of our product absorb and mentally process what we give them. That is the chemistry that ultimately determines the impact of the information collected. Shortfalls in strategic analysis and problems of cognitive bias are often entwined in dangerous combinations; the art and practice of intelligence must include a conscious and deliberate effort to disentangle and correct them.¹⁸¹

The end of this chapter will discuss how commitments to the production of strategic intelligence are, in fact, organizational investments in developing, exploring, and challenging mindsets. Therefore, these two conclusions are actually manifestations of the same cognitive phenomenon at individual and organizational levels.

Explaining intelligence failures on the basis of mindsets has not been without controversy. First, identifying the cause of failures as a result of innate cognitive processes has led to criticism regarding the deflection of responsibility for costly choices and errors in judgment.¹⁸² Second, because mindsets are grounded in cognitive processes, many believe that little can be done to mitigate their effects and influence over analysis.¹⁸³ However, as Jack Davis noted, the fact that cognitive errors cannot be eliminated does not mean they cannot be understood and compensated for, suggesting

¹⁸¹ Douglas MacEachin, “Analysis and Estimates,” in Jennifer E. Sims and Burton Gerber, eds., *Transforming U.S. Intelligence* (Washington, DC: Georgetown University Press, 2005), p. 117.

¹⁸² Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York, NY: Columbia University Press, 2008), p. 27.

¹⁸³ Jack Davis, “Why Bad Things Happen to Good Analysts,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 159.

that tradecraft can mitigate the negative effects of cognitive artifacts and that reasoned assessments of responsible and irresponsible behavior can be made.

In explaining Why Bad Things Happen to Good Analysts, cognitive biases—which are essentially unmotivated (i.e., psychologically based) distortions in information processing—have to be distinguished from motivated biases (distortions in information processing driven by worldview, ideology, or political preference). These cognitive biases cluster into the most commonly identified villain in postmortem assessments of intelligence failure: *mindset*.... Though there is no way of slaying this dragon, analysts can learn ways to live with it at reduced peril.¹⁸⁴

In spite of innate cognitive limitations posed by mindsets, several advancements in analytic processes have nevertheless been made. First, the intelligence community has turned to the theory of bounded rationality and cognitive science in an effort to better understand the kinds of biases that affect analysts and analytic production. This provided a richer understanding of how information and expertise are employed in analytic production. Second, investigations into mindsets have led to the recognition that analysts' expertise is simultaneously the cause of successes and failures, changing the ways in which analysts and analytic processes are evaluated.

One of the most important features of mindsets is that they occur in the presence of analytic expertise and experience. This realization has been important for shifting how intelligence failures are assessed. Rather than assume failures result from negligence, incompetence, or a lack of expertise, the consideration of mindsets and cognitive biases has redirected examinations to the ways in which expertise is employed and evaluated within the production and management of intelligence assessments.

¹⁸⁴ Jack Davis, "Why Bad Things Happen to Good Analysts," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 159-160.

In the intelligence world, a mindset usually represents “substantive expertise” and is akin to the academic concept of master of “normal theory”—judgments based on accumulated knowledge of past precedents, key players, and decision-making processes. Such expertise is sought after and prized. The CIA’s Directorate of Intelligence strategic plans invariably call for greater commitment of resources to in-depth research and more frequent tours of duty abroad for analysts—which amounts to building an expert’s mindset....

All analysts can fall prey to the perils of cognitive biases. A case can be made that the greater the individual and collective expertise on an issue, the greater the vulnerability to misjudging indicators of developments that depart from the experts’ sense of precedent or rational behavior. In a word, substantive experts have more to unlearn before accepting an exceptional condition or event as part of a development that could undermine their considerable investment in the dominant paradigm or mindset....

The “paradox of expertise” explains why the more analysts are invested in a well-developed mindset that helps them assess and anticipate normal developments, the more difficult it is for them to accept still inconclusive evidence of what they believe to be unlikely and exceptional developments.¹⁸⁵

As a result, techniques for coping with mindsets have shifted from the need to get “better” or more “expert” analysts, towards greater attention to managerial and methodological solutions that can better cope with the paradox that expertise and experience present. Most often, these remedies are simple (although not always practical), and tied directly to matters of organization and production processes that ensure specialists and generalists work together on assessments.¹⁸⁶

By viewing expertise as the source of mindsets, the cognitive challenges of intelligence are grounded in the theory of bounded rationality and the use of heuristics to

¹⁸⁵ Jack Davis, “Why Bad Things Happen to Good Analysts,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 160-161.

¹⁸⁶ Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York, NY: Columbia University Press, 2007), pp. 57-62.

cope with uncertainty and complexity. These heuristics, or biases in the context of intelligence analysis, include:

- *Anchoring*: the grounding of current assessments in the assumptions employed in prior assessments. Anchoring limits the extent to which prior frameworks and judgments are reevaluated. As a result, prior judgments anchor future judgments and limit the extent to which new information affects established analytic lines;¹⁸⁷
- *Availability*: the tendency to overweight dramatic, recent, and personal experiences when considering probabilities of potential futures. Availability substitutes one question for another by subtly changing the basis upon which estimates are made from the distribution of past occurrences to the ease by which examples come to mind;¹⁸⁸
- *Confirmation*: the tendency to search for or interpret information in a way that confirms the individual's preconceptions or preferences. Analysts will often seek out or give more weight to evidence that confirms a current hypothesis or the "conventional wisdom" while dismissing or devaluing disconfirming information;¹⁸⁹
- *Groupthink*: a concept that refers to faulty group decision making, and prevents the search for and consideration of alternatives in favor of achieving unanimity. Groupthink occurs when small groups are highly cohesive and must reach decisions under severe time pressures. Groupthink often compounds the presence of other cognitive errors made by individual analysts;¹⁹⁰
- *Hindsight*: the inclination to see past events as being more predictable than they appeared before they occurred. Analysts tend to remember their own past predictions as being more accurate than they were—becoming biased by knowledge of an outcome;¹⁹¹
- *Layering*: the use of judgments or assumptions made in one assessment as the basis for judgments in another. Layering occurs when judgments or

¹⁸⁷ Roger Z. George and James B. Bruce, "Glossary of Analysis Terms," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 310.

¹⁸⁸ Daniel Kahneman, *Thinking, Fast and Slow* (New York, NY: Farrar, Straus and Giroux, 2011), pp. 129-137.

¹⁸⁹ Roger Z. George and James B. Bruce, "Glossary of Analysis Terms," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 310.

¹⁹⁰ Roger Z. George and James B. Bruce, "Glossary of Analysis Terms," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 313.

¹⁹¹ Roger Z. George and James B. Bruce, "Glossary of Analysis Terms," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 313.

assumptions of one assessment are transferred to the next, while the context in which they were made is not. Layering tends to amplify the credibility of earlier judgments or assumptions by taking previously tentative positions and making them permanent fixtures of later assessments;¹⁹²

- *Mirror Imaging*: a cognitive error that occurs when analysts presume that another actor will interpret situations and behave as they would;¹⁹³
- *Nonregressive Inference*: this occurs when analysts fail to consider regression to the mean when considering targets' behavior or the occurrence of events. Failure to account for regression overestimates agency, discounts stochastic forces such as luck, and does not account for base rates when thinking about the actions and behavior of intelligence targets;¹⁹⁴
- *Rationality*: the assumption that actors perceive and evaluate situations without error and are capable of making optimal choices. Believing that targets are rational actors prevents analysts from considering when actors may take actions against their perceived interests, or paying attention to the decision-making processes that will influence what options are considered and how they are evaluated.¹⁹⁵
- *Substitution*: the tendency to replace a question that an analyst cannot answer with one that he or she can. Substitution allows analysts to reach conclusions about novel questions by reasoning from other cases or problems they have experience with until they believe they have arrived at an approximate solution to the question they seek to answer.¹⁹⁶

Each of these heuristics serve to contribute to the development and operation of analytic paradigms by characterizing how analysts frame problems, search for and interpret information, evaluate answers to questions, and determine their relative confidence in their assessments. These paradigms may be hard-fought lessons learned from experience,

¹⁹² Lowenthal, *Intelligence: From Secrets to Policy* (Washington, DC: Congressional Quarterly, 2009), p.121; and James B. Bruce, "Making Analysis More Reliable: Why Epistemology Matters to Intelligence," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 181.

¹⁹³ Roger Z. George and James B. Bruce, "Glossary of Analysis Terms," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 315.

¹⁹⁴ Daniel Kahneman, *Thinking, Fast and Slow* (New York, NY: Farrar, Straus and Giroux, 2011), pp. 175-184.

¹⁹⁵ Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York, NY: Columbia University Press, 2007), pp. 57-62; and James J. Wirtz, "The American Approach to Intelligence Studies," in Lock K. Johnson, ed., *Handbook of Intelligence Studies* (New York, NY: Taylor and Francis, 2009), p. 33.

¹⁹⁶ Daniel Kahneman, *Thinking, Fast and Slow* (New York, NY: Farrar, Straus and Giroux, 2011), pp. 97-104.

or the products of education, mentoring, and apprenticeship. In either case, they encapsulate normal theory in the context characterized by Kuhn, providing analysts with ready-to-use, shared assumptions about intelligence targets and problems that have proven successful in the past. However, as Betts noted, the normal circumstances for which normal theories are attuned may suddenly give way to exceptional, crisis events that are largely undertheorized in the social sciences and do not follow the paths prescribed by specialists' expertise and experience.

Since crises are rare, and war even more so, they represent aberrant behavior—that is, if the adversary handles international disputes diplomatically 99 percent of the time and forcibly only 1 percent, the forcible cases are deviations from an otherwise powerful theory. Thus, in contrast to long-range projections, which must rely heavily on explanations of past behavior, specific crisis-oriented predictions must be more concerned with exceptions to a powerful theory, with potentially deviant cases more than with typical ones, and with worst-case possibilities as well as best estimates of probability. Severe consequences of low probability threats take precedence over the higher probability of less intense threats. Whereas normal theory derives its power from categorical simplifications and parsimony, crisis predictions must dwell more on complexity, contingent propositions, and the residual risks within a usually accurate normal theory. In social science terms the second approach is almost atheoretical; I will call it exceptional thinking. The theory for this approach is waiting to be found since there is not yet an established, compelling formulation of how to go about predicting discontinuity. Exceptional thinking will include whatever ideas make the case for acting on the basis of improbable contingencies. For good reasons, this is always an uphill battle.¹⁹⁷

As a result, analysts may have little evidence to alert them to qualitative changes that may render their mindsets a hindrance to successful analysis rather than a help.¹⁹⁸

¹⁹⁷ Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York, NY: Columbia University Press, 2007), pp. 56-57.

¹⁹⁸ Jack Davis, "Why Bad Things Happen to Good Analysts," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 163.

By recognizing that the very expertise that enables analysts to provide credible support to policy makers provides the sources of analytic failure, efforts to cope with mindsets turned toward the character of bounded rationality and the importance of diversity. The link between bounded rationality, mindsets, and heuristics reveals why even when individuals have identical goals and information, they do not always interpret it in the same way or make the same decisions.¹⁹⁹ Instead, their underlying problem-solving approaches lead analysts to satisfice in their search for explanations for the behavior of targets and systems, meaning that some explanations are available to be discovered and evaluated while others remain out of reach. This insight served as the basis for Richards Heuer's work on analytic tradecraft and the development of SATs that started in the 1970s and continues through the present day. Heuer argued that satisficing produced analytic failures by affecting analysts' search processes—limiting the extent to which different interpretations of information were considered, and alternative hypotheses were considered and evaluated.

I would suggest, based on personal experience and discussions with analysts, that most analysis is conducted in a manner very similar to the satisficing mode (selecting the first identified alternative that appears “good enough”). The analyst identifies what appears to be the most likely hypothesis—that is, the tentative estimate, explanation, or description of the situation that appears most accurate. Data are collected and organized according to whether they support this tentative judgment, and the hypothesis is accepted if it seems to provide a reasonable fit to the data. The careful analyst will then make a quick review of other possible hypotheses and of evidence not accounted for by the preferred judgment to ensure that he or she has not overlooked some important consideration.

¹⁹⁹ Charles Perrow, *Complex Organizations: A Critical Essay* (New York, NY: McGraw Hill, 1986), p. 120-123; and Herbert A. Simon, *Administrative Behavior* (New York, NY: The Free Press, 1997), pp. 45-47.

This approach has three weaknesses: the selective perception that results from focus on a single hypothesis, failure to generate a complete set of competing hypotheses, and a focus on evidence that confirms rather than disconfirms hypotheses.²⁰⁰

Once analysts successfully paired perceptions of information with satisfactory explanations of events, they ceased to consider additional information or alternative frameworks and increased the opportunity for faulty judgments and surprise.

In order to cope with the prospects that analysts' experiences and expertise may mislead them, analytic tradecraft has introduced a variety of methods that are designed to challenge the individual and collective heuristics of analysts, originally referred to as Alternative Analysis (AA), now SATs. A small sampling of these techniques include Alternative Competing Hypotheses (ACH), Devil's Advocacy, Red Teaming, and What-If analysis, each of which are intended to overcome the natural tendency to satisfice.²⁰¹

Mindsets and Models

As the decade after the attacks of 9/11 and the 2002 NIE of Iraqi WMD came to a close, intelligence scholars and methodologists altered the term of art regarding the cognitive frameworks and worldviews employed by intelligence analysts. This change has been the replacement of the term "mindset" with "mental model," which has galvanized support for the development and use of AA and SATs, and changed the status

²⁰⁰ Richards J. Heuer, Jr., *The Psychology of Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 1999), p. 44.

²⁰¹ Richards J. Heuer, Jr., "Computer-Aided Analysis of Competing Hypotheses," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 256-257; and Timothy J. Smith, "Predictive Warning: Teams, Networks, and Scientific Method," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 269.

of mindsets from rigid cognitive defects into identifiable, improvable, and testable theories. For example, a 2011 book on SATs by Heuer and Randolph Pherson noted:

Discussions within the Intelligence Community and in the literature on intelligence analysis commonly refer to an analytic “mindset” as a principal source of analytic failure. A mindset “can be defined as the analyst’s mental model or paradigm of how government and group processes usually operate in country ‘X’ or on issue ‘Y.’” This is a key concept in understanding intelligence analysis. We, along with most analysts, have used the term “mental model” instead. The term “mindset” has a negative connotation for many people. It implies that if analysts were not so “set” in their ways, they would not make so many mistakes. But a mindset is neither good nor bad; it is unavoidable....

Why does it matter whether one uses the term “mindset” or “mental model”? It matters because it may affect how one tries to solve a problem. If an analyst’s mindset is seen as the problem, one tends to blame the analyst for being inflexible or outdated in his or her thinking. That may be valid in individual cases, as analysts do vary considerably in their ability and willingness to question their own thinking. However, if one recognizes that all analysis is based on fallible mental models of a complex and uncertain world, one might approach the problem differently. One might recognize that greater accuracy is best achieved through collaboration among analysts who bring diverse viewpoints to the table and the use of structured analytic techniques that assess alternative explanations or outcomes.²⁰²

By changing terms, the negative connotations of mindsets were replaced by discussions about how neutral, depersonalized mental models expressed the expertise, experience, and worldviews of analysts. By referring to mindsets as mental models, a degree of considered structure and logic is conferred to their holders, as well increased flexibility since it may only be possible to possess a single mindset but many alternative models.

The prospects for a model-centric tradecraft were first developed in the 1970s when then DCI William Colby encouraged the CIA’s Directorate of Intelligence (DI) to

²⁰² Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), pp. 5-6.

explore the application of formal modeling to political analysis after the behavioral revolution in the social sciences. The result was the creation of the Methods and Forecasting Division within the Office of Regional and Political Analysis.²⁰³ Heuer described its mission in terms of creating a more transparent set of analytic products:

Our task is not to try to replace the subjective wisdom of these specialists with so-called objective data, but to use rigorous methodological procedures to explicate and exploit more fully the insights and judgment of these analysts.²⁰⁴

Modeling techniques included the use of Bayesian analysis, cross-impact analysis, events data analysis, multivariate regression, and more.²⁰⁵ The division's emphasis was on exploring how quantitative modeling techniques could be adapted to study political problems that intelligence analysts faced, and identify the pragmatic tradeoffs between scholarly research and support to policy makers.

The challenge that was assigned to the Methods and Forecasting Division reopened older debates about the relationship between intelligence analysis and the social sciences, and the emphasis on prediction vs. exploration. Although numerous studies encouraged intelligence analysts to emulate the social sciences, whether their focus should be theoretical, empirical, or methodological was never settled. For example, Kent focused on incorporating historians into the analytic community, while Hilsman and

²⁰³ Richards J. Heuer, Jr., "Adapting Academic Methods and Models to Government Needs," in Richards J. Heuer, Jr., ed., *Quantitative Approaches to Political Intelligence: The CIA Experience* (Boulder, CO: Westview Press, 1978), pp. 1-2.

²⁰⁴ Richards J. Heuer, Jr., "Adapting Academic Methods and Models to Government Needs," in Richards J. Heuer, Jr., ed., *Quantitative Approaches to Political Intelligence: The CIA Experience* (Boulder, CO: Westview Press, 1978), p. 7.

²⁰⁵ Richards J. Heuer, Jr., ed., *Quantitative Approaches to Political Intelligence: The CIA Experience* (Boulder, CO: Westview Press, 1978).

Kendall argued that social scientists with strong theoretical skills would be more valuable.²⁰⁶ For example, Hilsman wrote:

In the first place, there is much room for improvement in the quality and composition of its staff; although there are able people in intelligence, more are needed. Also, there should be both a higher proportion of social scientists, and an on-the-job training program to help bring the rest up to a minimum standard. Then, too, the composition of the group of social scientists should be altered—at present, there is an unduly large group of historians in relation to men from the other disciplines.²⁰⁷

Kendall elaborated on Hilsman's view on the need for social science skills. He agreed with Kent on the ability of historians and other area specialists to provide consumers with high-quality basic and current intelligence because of their mastery over a region's geography, history, and cultural knowledge. However, Kendall argued that the most important strategic intelligence contributions would be made by social theorists.

It is a state of mind characterized by a crassly empirical conception of the research process in the social sciences. This, in view of the profound commitment of our intelligence agencies to what we have called the "regional breakdown" of their problem, is not surprising. For, if it is regional units you are building, and it is social scientists specialized to specific countries and areas you wish to staff them with, what you end up with is an extremely high percentage of historians, who with the best will in the world communicate to the operation the characteristic vices (and virtues) of their kind of research. The performance of the intelligence function accordingly becomes a matter of somehow *keeping* one's head above water in a tidal wave of documents, whose factual content must be "processed"—i.e., in Mr. Kent's language, "analyzed," "evaluated," and exploited as raw material for "hypotheses." The emphasis, as we have already noticed, is on prediction, which against this background is necessarily understood as a matter of projecting discernible empirical trends into an indefinite future.

²⁰⁶ Sherman Kent, *Strategic Intelligence for American World Policy* (Princeton, NJ: Princeton University Press, 1949); Willmoore Kendall, "The Function of Intelligence," *World Politics*, Vol. 1, No. 4 (July 1949), pp. 542-552; and Roger Hilsman, Jr., "Intelligence and Policy-Making in Foreign Affairs," *World Politics*, Vol. 5, No. 1 (October 1952), pp. 1-45.

²⁰⁷ Roger Hilsman, Jr., "Intelligence and Policy-Making in Foreign Affairs," *World Politics*, Vol. 5, No. 1 (October 1952), p. 41.

Here also the issue is fundamental: an intelligence operation built upon a conception of the research process in the social sciences that assigns due weight to “theory” as it is understood in economics and sociology and, increasingly one hopes, in politics, would of course be a wholly different affair. . . . This, clearly, is not the place to discuss the comparative merits of the two conflicting conceptions of social research, and it is not intended to suggest that the intelligence function should be delivered over entirely to that implied in the foregoing sentences. The point is, rather, that current planning and organization in the intelligence field ignores one of the two altogether. (It is, from this point of view, highly significant that Mr. Kent, for all his numerous references to the social sciences and social scientists, never employs in that connection the words *theory* and *theorist*.)²⁰⁸

While Kent, Kendall, and Hilsman had debated the relative balance and need for historians and more theoretically oriented social scientists, others questioned whether scholarly methods were applicable to intelligence assessments. At the same time the Methods and Forecasting Division was performing its initial efforts to apply formal models to intelligence analysis, Michael O’Leary, William Coplin, Howard Shapiro, and Dale Dean examined several hundred finished intelligence reports from the State Department’s Bureau of Intelligence and Research (INR) in order to assess whether formal models could replicate the analytic work being performed by mental models alone. Their study concluded:

...analyses found in the INR documents tend to be of the most demanding kinds, involving multivariate analyses with many discrete variables, in which the relationships are frequently nonlinear and involve important time lags. As a matter of fact, the kinds of relationships found in the great majority of INR analyses represent such complexity that no single quantitative work in the social sciences could even begin to test their validity.²⁰⁹

²⁰⁸ Willmoore Kendall, “The Function of Intelligence,” *World Politics*, Vol. 1, No. 4 (July 1949), pp. 550-551.

²⁰⁹ Michael K. O’Leary, William D. Coplin, Howard B. Shapiro, and Dale Dean, “The Quest for Relevance: Quantitative International Relations Research and Government Foreign Affairs Analysis,” *International Studies Quarterly*, Vol. 18, No. 2 (June 1974), p. 228.

Indeed, for the better part of the intelligence community's history, analytic methods have largely diverged from scholarly ones. Rather than replicate the ways models are used in the social sciences, the limited uses of formal modeling in the intelligence community have instead examined how they can assist in the treatment of mental models as discussed earlier. Thus, the content of social science and intelligence differ significantly, even if the overall logic of discovery and relationships between evidence and inference remain consistent with the scientific method.²¹⁰

For all the difficulties of applying rigorous methods, the Methods and Forecasting Division's efforts identified many important contributions that formal, mathematical modeling could provide to analysts. The most important contribution of formal models was the assistance in coping with the cognitive challenges of analytic assessment, a conclusion that set the groundwork for the development of SATs several decades later. Thus, intelligence methodologists quickly focused on the act of modeling as a way to improve analysis, rather than the construction and performance of formal models as independent artifacts as practiced in academic settings.

For the better part of the intelligence community's history, analytic methods have largely diverged from scholarship. While formal models have occasionally been employed in niche applications, e.g. the use of operations research to examine foreign military capabilities, they have not approached the level or diversity of uses found in the

²¹⁰ This point was repeated by Ben-Israel and his extensive examination of the philosophy of science in intelligence estimates. Specifically, he noted that it was essential to differentiate between employing the content of scientific theories and models in intelligence, which he believed to be inappropriate, and replicating the scientific method in analytic tradecraft, which he argued was essential to improve analysis. Isaac Ben-Israel, "Philosophy and Methodology of Intelligence: The Logic of Estimate Process," *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 689.

scholarly community. More importantly, the intelligence community identified the process and ancillary effects of formal modeling as the primary benefits to analytic tradecraft and the exposition and exploration of mental models, rather than model results themselves.

As long as intelligence research is directed towards answering complex questions such as what will happen in Yugoslavia after Tito's death, or what would be the consequences of Communist party participation in the Italian government, the narrative essay will remain the dominant form for intelligence estimates.

There is, however, an important role for rigorous procedures even in such complex estimative problems. Our work to date indicates that the kinds of analytical techniques which seem most useful for our purposes are those that help to trace the logical consequences of subjective judgments, extend the mental capacity of the individual analyst, force the analyst to make his assumptions explicit, or help to organize complexity.²¹¹

By the 1990s, Heuer, Davis, and others had developed analytic techniques intended to expose, examine, and communicate mental models; aid analysts in coping with ambiguous or sparse data; and institute a hypothesis-driven tradecraft that sought to focus collection and analysis on diagnostic elements of information in order to eliminate alternative explanations of events. The introduction of ACH and other techniques originally developed under tradecraft initiatives such as Tradecraft 2000, AA, and most recently SATs all sought to add rigor and transparency to analytic assessment within the context of the predominantly intuitive and literary culture of the intelligence

²¹¹ Richards J. Heuer, Jr., "Adapting Academic Methods and Models to Government Needs," in Richards J. Heuer, Jr., *Quantitative Approaches to Political Intelligence: The CIA Experience* (Boulder, CO: Westview Press, 1978), p. 8.

community.^{212, 213}

Implementing Structured Analytic Techniques

The notion that analysts develop, assess, and communicate the design, structure, and outputs of their mental models is not firmly institutionalized within the intelligence community. The tendency is for analysts to view their contributions to policy makers in terms of their judgments, and when expressed independently from the evidence and logic that support them, this undermines the notion of corporate production and transparency.²¹⁴ Moreover, many consumers remain epistemologically committed to a deterministic world, and therefore have little tolerance or patience for accepting the very uncertainties that SATs and modeling implicitly accept. While rigor and transparency can improve the practice of analysis, even perfectly executed tradecraft cannot guarantee that producers and consumers will not be surprised by events as they unfold.

[W]e need to remember that intelligence analysis is an intellectual process. It needs standards and guidelines but these alone will not ensure analysis that will produce the “right” answer. Indeed, there is no way to ensure the “right” answer.²¹⁵

²¹² Douglas J. MacEachin, “The Tradecraft of Analysis,” in Roy Godson, Ernest R. May, and Gary Schmitt, eds., *U.S. Intelligence at the Crossroads Agendas for Reform* (Washington, DC: Brassey’s, 1995), pp. 72-74; Richards J. Heuer, Jr., *The Psychology of Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 1999); Roger Z. George, “Fixing the Problem of Analytical Mind-Sets,” *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), pp. 385-404; and Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), pp. 8-9.

²¹³ The distinct literary culture of the intelligence community, particularly amongst the CIA’s DI that develops strategic intelligence on political, economic, military, and social issues was noted by John Hanley, the former Director of Strategy for the Office of the Director of National Intelligence. Interview with John Hanley, Office of the Director of National Intelligence, February 9, 2012.

²¹⁴ Interview with Joseph Eash III, National Defense University, September 6, 2012.

²¹⁵ Mark M. Lowenthal, “Towards a Reasonable Standard for Analysis: How Right, How Often on Which Issues?” *Intelligence and National Security*, Vol. 23, No. 3 (June 2008), pp. 313-314.

Tradecraft should be regarded as a metaheuristic for ensuring that the particular mental models of analysts are well specified, and constantly pressed to look beyond their prior experience and check against numerous cognitive biases that affect intelligence analysis. It is not an algorithm for predicting the future or guaranteeing that surprises never occur.

The journey from the Methods and Forecasting Division to contemporary SATs was an uneven one, where major motivations to improve tradecraft were evolutionary adaptations made in response to analytic failures. A chain of events, from the surprising conclusion of the Cold War and disintegration of the Soviet Union through the twin failures of 9/11 and the 2002 Iraq WMD estimate created internal and external pressures for reforming tradecraft. In each case, SATs or their methodological predecessors, Tradecraft 2000 or AA, were cited as crucial additions to analytic practices that warranted increased usage by analysts.²¹⁶

Multiple commissions in the 1990s concluded that greater emphasis to red team analysis, one of the techniques under the SAT umbrella where analysts seek to understand the world and US through the eyes of intelligence targets, could reduce the vulnerability of policy makers to surprise and enhance their confidence in the judgment and expertise of analysts.

...after Adm. David Jeremiah's postmortem analysis of the Intelligence Community's failure to foresee India's 1998 nuclear test, a U.S. congressional commission's review of the Intelligence Community's

²¹⁶ Douglas J. MacEachin, "The Tradecraft of Analysis," in Roy Godson, Ernest R. May, and Gary Schmitt, eds., *U.S. Intelligence at the Crossroads Agendas for Reform* (Washington, DC: Brassey's, 1995), pp. 72-74; Richards J. Heuer, Jr., *The Psychology of Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 1999); Roger Z. George, "Fixing the Problem of Analytical Mind-Sets," *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), pp. 385-404; and Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), pp. 8-9.

global missile forecast in 1998, and a report from the CIA Inspector General that focused higher level attention on the state of the Directorate of Intelligence's analytic tradecraft. The Jeremiah report specifically encouraged increased use of what it called "red team" analysis.²¹⁷

A decade later, following AQ's attacks on September 11, 2001 and the 2002 NIE on Iraq's WMD capabilities, an even stronger emphasis was placed on analytic tradecraft and the need to strengthen analytic transparency, rigor, and clarity. In reflecting upon the Iraqi WMD NIE's tradecraft, the investigating commission noted:

Analytic "tradecraft"—the way analysts think, research, evaluate evidence, write, and communicate—must be strengthened. In many instances, we found finished intelligence that was loosely reasoned, ill-supported, and poorly communicated. Perhaps most worrisome, we found too many analytic products that obscured how little the Intelligence Community actually knew about an issue and how much their conclusions rested on inference and assumptions. We believe these tendencies must be reversed if decisionmakers are to have confidence in the intelligence they receive. And equally important, analysts must be willing to admit what they don't know in order to focus future collection efforts. Conversely, policymakers must be prepared to accept uncertainties and qualifications in intelligence judgments and not expect greater precision than the evaluated data permits.²¹⁸

The commission went on to identify the importance of AA, now referred to as SATs, and the need for an institutional commitment to their use.

Mission Managers should not be responsible for providing a single, homogenized analytic product to decisionmakers; rather, Mission Managers should be responsible for encouraging alternative analysis and for ensuring that dissenting views are expressed to intelligence customers. In sum, Mission Managers should be able to find the right people and

²¹⁷ Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), p. 9. Also see Roger Z. George, "Fixing the Problem of Analytical Mind-Sets: Alternative Analysis," *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), pp. 391-393.

²¹⁸ Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction, *Report to the President of the United States* (Washington, DC: Government Printing Office, March 31, 2005), pp. 12-13.

expertise and make sure that the right analysis, including alternative analysis, is getting done.²¹⁹

Additional assessments by the CIA's DI similarly concluded that the full range of available techniques was underutilized by analysts, and that greater use of SATs would have improved analytic performance.

In theory, use of alternative analysis techniques can help to reduce the likelihood of "intelligence failures," which historically have stemmed in part from such mental errors (e.g. the ingrained belief that the Japanese could not mount a successful attack against Pearl Harbor). In reality, however, alternative analysis has not been particularly effective within the Intelligence Community as it has been employed only sporadically, at best, and more often than not as a "nice to have" supplement tacked on to traditional analysis rather than integrated at the outset as an essential component of the analytic enterprise in a world of uncertainty and deception.²²⁰

Despite the pressure exerted by internal and external reviews, SATs have remained underutilized and employed primarily as an adjunct to more traditional analytic approaches. Their limited adoption has been largely attributed to two factors. First, the term "alternative analysis," which served as the term of art during the early 2000s, suggested to analysts that such techniques were special and not necessary for mainstream, regular analytic production.

Over time, however, analysts who misunderstood or resisted this approach came to interpret alternative analysis as simply meaning an alternative to the normal way that analysis is done, implying that these alternative procedures are needed only occasionally in exceptional circumstances when an analysis is of critical importance. Kent School instructors had to explain that the techniques are not alternative to traditional analysis, but that they are central to good analysis and should be integrated into the

²¹⁹ Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction, *Report to the President of the United States* (Washington, DC: Government Printing Office, March 31, 2005), p. 25.

²²⁰ Kent Center, "Rethinking 'Alternative Analysis' to Address Transnational Threats," *Occasional Papers*, Vol. 3, No. 2 (October 2004), p. 1.

normal routine—instilling rigor and structure into the analysts’ everyday work process.²²¹

Second, production schedules often compelled analysts to streamline their practice of tradecraft, cutting out activities that were time consuming, while anchoring or layering assumptions and judgments from prior products in order to give nominal emphasis to current events. These practices limited the extent to which the benefits of SATs were reaped in the production of strategic intelligence while biasing analytic attention and effort toward current intelligence. For example, reviews of the 2002 Iraqi WMD NIE noted problems posed by the intertextual character of assessments resulted when the use of conclusions reached in prior analytic products were employed as assumptions or evidence upon which new judgments were made in newer products.²²²

Habit of thought was effectively admitted by the CIA’s deputy director of intelligence (the agency’s senior manager of analysts) in a speech to analysts, in which she faulted the practice of “inherited assumptions” that went unquestioned in the NIE process.²²³

This problem of layering was so acute that it became one of the major targets of intelligence reform and tradecraft improvements instituted by the creation of the ODNI. As Fingar noted when comparing the 2002 NIE of Iraq’s WMD and later estimates of Iran’s WMD capabilities, the requirement to start a zero-based review was critical to

²²¹ Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), p. 9.

²²² James Der Darian has argued that intelligence analysis is essentially hermeneutic in character, where assumptions, models, and judgments exist in a network of interconnected texts. James Der Darian, “Anti-Diplomacy, Intelligence Theory and Surveillance Practice,” *Intelligence and National Security*, Vol. 8, No. 3 (July 1993), pp. 37-39.

²²³ James B. Bruce, “Making Analysis More Reliable: Why Epistemology Matters to Intelligence,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 181.

ensure that each assumption and piece of a data was reevaluated in the production of the NIE.

Specifying the key questions to be addressed defines the scope and focus of the estimate, but that is only the first step in a multistage process designed to produce a genuinely fresh look at old and new evidence, previous judgments, and alternative assumptions and hypotheses. In other words, no matter how many times or how recently the IC has examined the issues to be addressed in an estimate, the process mandates going back to basics and starting with a clean sheet of paper. That is what is supposed to happen, and to a large extent it does. Even though the subjects addressed by all estimates have been examined—and written on—by analysts in one or (usually) more elements of the IC, and both analysts and agencies may well have adopted or endorsed positions on one or more of the questions to be examined, the preparation of an estimate requires, to the extent possible, a zero-based analysis of both previous analytic judgments and the intelligence used to support them. This means going back to square one in some very specific ways. Previous judgments, especially if they are deemed germane to answering the central question or any of the prior questions to be addressed in the estimate, must be examined *de novo*; no judgment can be assumed to be still valid and simply carried forward as the starting point for subsequent analysis. This was one of the clearest and most important lessons learned by the analytic community and others who examined the flawed Iraq WMD estimate.²²⁴

The widespread causes and consequences of layering demonstrated how many of the practical problems of intelligence analysis can induce failure. Organizational and psychological pressure to build on previous analytic products limits the extent to which new interpretations of the available information are considered. Moreover, the desire to produce findings that are consistent with earlier products militates against examining each piece of information and assumption as frequently and vigorously as recommended. Therefore, from an operational perspective, analysts often know what they should do, as

²²⁴ Thomas Fingar, *Reducing Uncertainty: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), p. 82.

Fingar noted in discussing the importance of zero-based analysis, but practical concerns over production time and consistency limit the implementation of proper tradecraft.

Another barrier to the adoption of SATs has been the producer/consumer relationship itself. In cases where decision makers believe that intelligence analysis should reduce their uncertainties and “make the call,” analytic products that expose consumers to disagreements between analysts and present the full complexity and indeterminacy of intelligence problems can be interpreted as avoiding taking responsibility for analytic judgments. SATs and the increased emphasis on presenting the extent and character of key uncertainties to consumers have often resulted in products that are interpreted as deciding that anything can happen or making overly vague claims that cannot be precisely interpreted. Ben-Israel noted this defensive tendency on the part of intelligence producers and its detrimental effects on consumers.

The “cost of error” in intelligence estimate is very high at both individual and substantial level (the recommendations of the Agranat Committee serve as a suitable example). A wrong intelligence estimate can lead to grand disaster (e.g., the Yom-Kippur War). Intelligence officers, therefore, tend to be over-careful, to give conventional estimates, to avoid raising bold conjectures and to use vague language keeping all options open.

The “shortage of time” available to intelligence officers, and the pressure exerted on them to give immediate answers and estimates, have the same effect: they drive intelligence officers to give conventional risk-free estimates, without committing themselves to any unambiguous position....

An uncommitted opinion or truism is irrefutable, and thus considered “safer.” (For whom? For the estimate consumer or for the estimator?) But there is a price to pay: our knowledge will not increase.²²⁵

²²⁵ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 670.

While intelligence analysis should provide policy makers with clear statements about the international system, well-performed analysis may be just as likely to increase the uncertainty of policy makers as reduce it. Thus, the emphasis of intelligence is not on precisely calculating the likelihood of events, but on properly characterizing the state of the community's knowledge about the world and identifying the robustness of their assessments.²²⁶ Fingar noted that informing policy makers of when they should act with confidence or recognize the extent to which their beliefs were likely to change was often more important than the specific content of particular assessments.

...if, at the end of the process, analysts continue to reach different judgments after considering the same information, that fact is as important to convey to policy makers as are the contending judgments themselves. Simply stated, if smart analysts with access to the same information come to different conclusions after thorough vetting of the intelligence and transparent discussion of how they reached their judgments, neither they nor policy makers can have high confidence in either/any of the contested judgments. Conveying the warning that the ice under judgments is thin is or can be even more important than the judgments themselves.²²⁷

This point was demonstrated during the assessment of intelligence regarding the occupant of the compound in Abbottabad Pakistan that was raided by US Navy SEALs resulting in the killing of Osama bin Laden (OBL). In an effort to ensure that all possible hypotheses were considered, multiple SATs, particularly Red Team analyses, were employed to challenge analysts' judgments regarding the identity of its unknown occupant. Former CIA analyst and manager John Brennan, then serving in the White House, asked analysts to marshal the available evidence against their judgments in an effort to test the strength

²²⁶ Interview with Leon Fuerth, National Defense University, September 21, 2011.

²²⁷ Thomas Fingar, *Reducing Uncertainty: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), p. 84.

of their inferences and ensure that they had not overlooked any information that might have lowered their confidence that they had found OBL.

John Brennan, the longtime CIA officer who was now Obama's top counterterrorism advisor, met regularly with the analysts working the bin Laden case, many of whom he had known and admired for years. Brennan pushed them to come up with intelligence that disproved the notion that bin Laden was living in the Abbottabad compound, saying, "I'm tired of hearing why everything you see confirms your case. What we need to look for are the things that tell us what's not right about our theory. So what's not right about your inferences?"²²⁸

An important result of these analytic exercises, particularly the employment of red team exercises, was a reduction in analytic confidence in the conclusion that OBL was the mysterious occupant of the Abbottabad compound.

Ben Rhodes says, "There was a deflation in the room, because what you're looking for as you're getting closer to the call is greater certainty, not less. So essentially it played into all the fears that people had about what could go wrong. Is it worth the risk?"

Similarly, Tony Blinken says, "I think, if anything, the Red Team actually brought down the level of certainty; the positive ID percentage was higher before the Red Team got done. So I think we went from maybe seventy/thirty or sixty-five/thirty-five to fifty-five/forty-five or even fifty/fifty."²²⁹

Controversies have occurred over the use of SATs, particularly when analytic pieces intended to challenge analytical assumptions, advance contrarian positions, and present an expanded range of "what-if" scenarios to consumers. Apparent inconsistencies between mainline analytic judgments, punctuated by alternative perspectives, can inadvertently undermine consumers' confidence in the analytic community by sending mixed messages about analysts' expectations and assessments.

²²⁸ Peter L. Bergen, *Manhunt: The Ten-Year Search for Bin-Laden: From 9/11 to Abbottabad* (New York, NY: Crown, 2012), p. 134.

²²⁹ Peter L. Bergen, *Manhunt: The Ten-Year Search for Bin-Laden: From 9/11 to Abbottabad* (New York, NY: Crown, 2012), p. 197.

[S]ome policymakers continue to prefer precise, single-point judgments from intelligence analysts; to these decisionmakers, the DIs resorting to “what if” or contrarian analysis undermines their confidence in what the CIA or other IC analysts are telling them. If on Monday, the CIA asserts that an Indian nuclear test is unlikely, but on Tuesday writes an analysis suggesting—in the absence of compelling information—that New Delhi might possibly see things differently and could be planning to test secretly, then the reader might not know what the Agency thinks is most likely. Unless [SATs] are carefully explained and used to sensitize policymakers to important uncertainties in the CIA’s analysis, the risk exists of its being considered “CYA” analysis.²³⁰

SATs have also created professional and organizational challenges within the analytic community itself. The natural evaluation and assessment of alternative mental models and explorations of worst-case or “what-if” scenarios has led to charges of politicization, where intellectual exercises become fodder for policy makers searching for analytic products that support their favored perspective or conclusions.²³¹

Alternatively, analysts have often viewed their judgments as the value they provide to consumers, and have often resisted efforts to make the data and logic of their reasoning more transparent. From this perspective, SATs that encourage the explication of analysts’ use of evidence and inference are seen as devaluing analysts by providing consumers opportunities to substitute their own judgments.

Some DI analysts also have expressed concern that laying all of this out in their products will lead consumers to do their own analyses. Most consumers, particularly those directly engaged on a policy issue, are already doing that with or without the intelligence product. They have their own channels of information, often they have much of the same raw intelligence possessed by the analyst, and they often know the main foreign players from direct interaction.

²³⁰ Roger Z. George, “Fixing the Problem of Analytical Mind-Sets: Alternative Analysis,” *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), p. 399.

²³¹ Roger Zane George, “Beyond Analytic Tradecraft,” *International Journal of Intelligence and Counterintelligence*, Vol. 223, No. 2 (2010), pp. 301-303.

Consumers, present and past, have consistently stated that for them, the value added—and the credibility—of the intelligence product is directly dependent on the information conveyed, its reliability, and their understanding of the analytic logic that supports the conclusions. If these are not made explicit and clear, the intelligence product becomes simply an opinion that may be agreed with or swept aside.²³²

The widespread use of SATs within the intelligence community remains an unfulfilled goal. Although these techniques assist analysts in many of their individual and collective tasks, such as organizing and visualizing data, sharing information with peers, and otherwise ensuring that a wider range of relationships or dependencies that characterize the behavior and decision making of increasingly complex targets receive consideration, three reasons continue to limit their use. First, the techniques available to analysts are rapidly proliferating, increasing the demand for training and occasionally a dependence on in-house methodologists, specialists from other directorates, or outside academic or contractor expertise. For example, the CIA's 2009 publication *A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis* provided brief summaries on a dozen techniques. Two years later, Heuer and Pherson's *Structured Analytic Techniques for Intelligence Analysis* described more than fifty.²³³ Second, the demands on analysts to address an increasingly diverse array of consumer needs in a rapid fashion limit the extent to which they can practice tradecraft in its most idealized fashion. Finally, striking the appropriate balance between total transparency and opacity

²³² Douglas J. MacEachin, "The Tradecraft of Analysis," in Roy Godson, Ernest R. May, and Gary Schmitt, eds., *U.S. Intelligence at the Crossroads Agendas for Reform* (Washington, DC: Brassey's, 1995), p. 70.

²³³ Central Intelligence Agency, *A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 2009); and Richards J. Heuer, Jr. and Randolph Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011).

in the producer/consumer relationship remains challenging and largely contingent on the specific characteristics of the intelligence problem, analyst, and policy maker in question.

Mirror-Imaging and Substitution in the Assessment of Iraqi WMD

Before discussing the specific applications of SATs it is important to look at two specific ways in which mindsets or cognitive biases and analytic heuristics affected US intelligence assessments and policy analysis on Iraqi WMD. As noted in the prior section, the failure to implement proper tradecraft was evident in the 2002 NIE on Iraqi WMD. However, most criticisms identified above addressed layering and anchoring, two heuristics that cannot produce analytic failures unless the assumptions and conclusions they propagate are flawed. Moreover, because NIEs are products that summarize the state of intelligence collection and assessments of particular subjects, they rarely advance judgments that cannot be found in previous intelligence assessments.²³⁴ Therefore, it is important to consider two additional sources of failure that directly produced inferential errors: mirror-imaging and substitution.

Mirror-imaging, and, implicitly, appeals to rational choice, were noted in an internal study of analytic assessments performed by the CIA regarding how analysts misunderstood the behavior of Saddam Hussein's regime. In *Misreading Intentions: Iraq's Reaction to Inspections Created Picture of Deception*, the DI determined that analysts had falsely assumed that US national security priorities were the same as Iraqi

²³⁴ Thomas Fingar, *Reducing Uncertainty: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), pp. 67-87.

concerns, and therefore engaged in mirror-imaging when attempting to understand the regime's behavior.

Analysts tended to focus on what was most important to us—the hunt for WMD—and less on what would be most important for a paranoid dictatorship to protect. Viewed through an Iraqi prism, their reputation, their security, their overall technological capabilities, and their status needed to be preserved. Deceptions were perpetual and detected, but the reasons for those deceptions were misread....

Analysts understood that the Iraqis were working with a different logic system, but did not go far enough in accounting for how greatly Iraqi and Western thought differs.²³⁵

By assuming that Iraqi and the US leadership viewed the threat of WMD and the importance of inspections in the same way, analysts believed that the regime's behaviors were indicative of continued defiance of UN resolutions and deception. In that context, the regime's destruction of materials and documents were seen as destroying evidence to thwart inspectors, rather than as acts of unilateral disarmament.²³⁶

A second source of failure regarding Iraqi WMD resulted from substitution, where general patterns of voluntary disarmament and knowledge of specific cases concerning South Africa, Ukraine, and Kazakhstan were used. In this case, directly answering questions about Iraq's WMD capabilities was not possible given limited intelligence, much of which was faulty, so analysts and policy makers substituted questions about Iraq with others they could. The specific experiences of three states that had voluntarily dismantled their nuclear weapons programs served as a temple for

²³⁵ Central Intelligence Agency, *Misreading Intentions: Iraq's Reaction to Inspections Created Picture of Deception* (Washington, DC: Central Intelligence Agency, January 5, 2005), p. 16 (declassified on June 5, 2012).

²³⁶ Central Intelligence Agency, *Misreading Intentions: Iraq's Reaction to Inspections Created Picture of Deception* (Washington, DC: Central Intelligence Agency, January 5, 2005), pp. 10-14 (declassified on June 5, 2012).

making general inferences about Iraq. This enabled inferences to be made about Iraq based on the particular experiences of other states. This analytic process culminated in a *New York Times* op-ed by Condoleezza Rice that argued Iraq must be lying about its WMD capabilities because it did not behave like South Africa, Ukraine, and Kazakhstan.

There is no mystery to voluntary disarmament. Countries that decide to disarm lead inspectors to weapons and production sites, answer questions before they are asked, state publicly and often the intention to disarm and urge their citizens to cooperate. The world knows from examples set by South Africa, Ukraine and Kazakhstan what it looks like when a government decides that it will cooperatively give up its weapons of mass destruction. The critical common elements of these efforts include a high-level political commitment to disarm, national initiatives to dismantle weapons programs, and full cooperation and transparency....

Iraq's behavior could not offer a starker contrast....

Unlike other nations that have voluntarily disarmed—and in defiance of Resolution 1441—Iraq is not allowing inspectors “immediate, unimpeded, unrestricted access” to facilities and people involved in its weapons program. As a recent inspection at the home of an Iraqi nuclear scientist demonstrated, and other sources confirm, material and documents are still being moved around in farcical shell games. The regime has blocked free and unrestricted use of aerial reconnaissance.^{237, 238}

The problems of mirror-imaging and substitution suggest the limits of SATs.

Heuer noted that had the particular hypothesis that Iraq's WMD programs were dormant been included in an ACH analysis, analysts and policy makers would have had to confront the relative weakness of the information upon which their assessment of Iraq's programs rested. Although analysts were unlikely to have honed in on the dormant

²³⁷ Condoleezza Rice, “Why We Know Iraq is Lying,” *The New York Times*, January 23, 2003, <http://www.nytimes.com/2003/01/23/opinion/why-we-know-iraq-is-lying.html?pagewanted=all&src=pm> (accessed on March 22, 2012).

²³⁸ Technically, Rice's argument was not an intelligence assessment. However, her argument rested on creating an argumentative framework that replicated analytic processes within the intelligence community and employed intelligence assessments that described how the Iraqi regime was behaving.

hypothesis as the strongest explanation of Saddam Hussein's behavior and capabilities, its robustness to disconfirming evidence would have been revealing its own right.

It is unlikely that ACH analysis of the Iraqi nuclear program prior to the 2003 invasion of Iraq would have arrived at the correct conclusion: that the program was dormant. This is because of the history of Saddam Hussein's determination to obtain and willingness to use weapons of mass destruction (WMD), and strong memories of discovering after the 1991 Iraq war that intelligence estimates had underestimated Iraq's inventory of WMD. Analysts were determined not to underestimate the Iraqi WMD again.

However, an ACH analysis would have at least placed an important hypothesis on the table—that "Iraq is not now trying to rebuild its nuclear weapons program." If that hypothesis had been considered, it might have been unexpectedly difficult to refute.²³⁹

Mirror-imaging and substitution are arguably more consequential than layering and anchoring because they are direct products of mental models, not organizational processes. Instead, they reflect the ways in which analysts framed problems and evaluated information. Errors of these types are where SATs in general and ABM more specifically can provide assistance to analysts and limit the extent to which analytic assessments are not merely artifacts of their assumptions.

Structured Analytic Techniques and the Manual Method

SATs have traditionally been divided into three broad categories, although there is a high degree of overlap between them. These three categories are:

- Diagnostic techniques;
- Contrarian techniques;
- Imaginative thinking techniques.²⁴⁰

²³⁹ Richards J. Heuer, Jr., "Computer-Aided Analysis of Competing Hypotheses," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 261-262.

²⁴⁰ Central Intelligence Agency, *Structured Analytic Techniques: A Primer* (Washington, DC: Central Intelligence Agency, 2009), p. 5.

Each of these broad categories contains specific methods, such as Key Assumptions Check, Quality of Information Check, Indicators or Signposts of Change, ACH, Devil’s Advocacy, Team A/Team B, High-Impact/Low-Probability Analysis, “What-If” Analysis, Brainstorming, Outside-In Thinking, Red Team Analysis, Alternative Futures Analysis (Scenarios), and more.²⁴¹ Viewed individually, each technique focuses on a particular act of the analytic process, from identifying and articulating assumptions to challenging analysts to envision alternative worlds to estimating the likely values or content of missing information. When viewed collectively, the universe of SATs reveals that all of them are manual methods—able to be performed with pen and paper or simple spreadsheet software for making and organizing lists, tables, and matrices.²⁴²

The emphasis on manual techniques is no accident, as tradecraft developers within the analytic community have been acutely aware that the limited time available to analysts largely precludes cumbersome quantitative and computational modeling efforts

²⁴¹ The specific techniques are briefly described in Central Intelligence Agency, *Structured Analytic Techniques: A Primer* (Washington, DC: Central Intelligence Agency, 2009). Additional techniques are discussed in greater detail in Don McDowell, *Strategic Intelligence: A Handbook for Practitioners, Managers, and Users* (Lanham, MD: Scarecrow Press, 2009); Hank Prunckun, *Handbook of Scientific Methods of Inquiry for Intelligence Analysis* (Lanham, MD: Scarecrow Press, 2010), pp. 135-161; Timothy Walton, *Challenges in Intelligence Analysis: Lessons from 1300 BCE to the Present* (New York, NY: Cambridge University Press, 2010); and Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011).

²⁴² Many of these tools have become computer based, allowing for increased collaboration in the performance of analytic tasks and enabling logical operations to be performed easily, such as generating tables of all possible combinations of interactions between variables. These efforts employ computation in the sense that software is aiding analysts in the execution of tradecraft, but there is no effort to use computers to develop or employ dynamic models in determining the consistency of evidence with alternative hypotheses. See Richards J. Heuer, Jr., “Computer-Aided Analysis of Competing Hypotheses,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 251-265; and Pherson Associates, “Analytic Tools and Techniques” at <http://www.pherson.org/Tools.html> (accessed on June 16, 2012).

in support of specific products.²⁴³ Indeed, on matters of current intelligence where the delivery of timely information places intelligence analysts in direct competition with open-source journalists and consumers' information networks, temporal limitations are likely to affect tradecraft and foreclose certain tradecraft options.

The temporal conditions and the opportunities to practice a more robust analytic tradecraft change considerably when dealing with strategic intelligence. When analysts have the opportunity to look far into the future beyond the current attention of policy makers and deal with enduring and persistent challenges in the international system, the limitations of these manual methods become apparent and demand for computational tools increases. As Ben-Israel noted with respect to time and strategic intelligence:

Grand intelligence problems...usually have an easy time-table. Emergencies, as when a suspicious aerial target is detected by radar, must be met by a quick "intelligence estimate": is it an enemy aircraft on the attack? Or has it made a navigation error? Or is it a false signal? Time here is short and a decision must be made intuitively. But when we deal with questions such as "What will happen in Syria after Assad?" we have all the time we need.²⁴⁴

Likewise, certain intelligence issues may be of such consequence with respect to the decisions they inform that producers and consumers may insist on multiple, independent assessments and employment of all available tradecraft techniques, as was demonstrated by the analysis of OBL's compound in Pakistan discussed earlier.

The weaknesses of an all-manual suite of SATs are apparent once intelligence problems are of sufficient complexity as to press against the cognitive limits of the

²⁴³ Sarah Miller Beebe, *A New Intelligence Technique for Analyzing Complex Adaptive Systems: A Case-Based Assessment*, International Studies Association Annual Conference, March 2011.

²⁴⁴ Isaac Ben-Israel, "Philosophy and Methodology of Intelligence: The Logic of Estimate Process," *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 695-696.

human brain. Although SATs may allow for the explication of analysts' mental models, they cannot animate these models through time or explore all possible combinations of variables or conditions without falling back on error-prone simplifications and generalizations. Thus, manual methods may help analysts understand the complexity of given intelligence problems, but may nevertheless be hamstrung by problems of scale, interdependence, feedback, and more. In order to illustrate the limitations of manual methods, two SATs, Alternative Futures Analysis and ACH, are examined in greater detail below.

Alternative Futures Analysis

Alternative Futures Analysis, more commonly known as scenario planning, has been a mainstay of the policy community for several decades. As a formal analytic methodology, their first major impact on executive decision making occurred in the commercial sector during the 1973 Yom Kippur War. In 1967, prior to the war, planners at Royal Dutch Shell started to work on scenarios in to order challenge senior executives to think about the international system and global energy markets decades ahead, rather than within their normal six-year planning horizon. Several scenarios were developed that characterized different alternative futures for the year 2000, which was chosen to account for the fact that classical forecasting techniques could not generate major discontinuities symbolizing qualitative transformations consistent with the narratives of social, economic, and technological changes of the time. The potential consequences of these changes drew the attention of Shell's internal planning staff who sought to get corporate leaders to consider a wide range of alternative futures in their decision making.

The exercise and challenge of those scenarios has been repeatedly credited with preparing Shell's executives to capitalize on the sixfold increase in oil prices that occurred as a result of the Yom Kippur War in 1973, and again in 1981 with the outbreak of the Iran–Iraq War.²⁴⁵

Scenarios have remained popular for planning and strategy in the business community, with the private sector leading the way in their methodological development and application.²⁴⁶ Within the government, the Office of Net Assessment (ONA) in the DOD, directed by Andrew Marshall, has been one of the leading users of scenarios to explore the future of warfare and the dynamics of military balances.

One of Marshall's principal tools for preparing for future threats while they are still dark clouds on the distant horizon rather than storms that are already upon it is the use of scenarios—stories about how future events might come to pass. Such scenarios can be used to help the Pentagon make better decisions in an uncertain world.²⁴⁷

Scenarios are also an essential part of military planning, encapsulated in preliminary war gaming and the development of branches and sequels that accompany operational plans in order to cope with potential contingencies.²⁴⁸

²⁴⁵ Pierre Wack, "Scenarios: Uncharted Waters Ahead," *Harvard Business Review*, No. 85516 (September-October 1985), pp. 73-89; and Thomas J. Chermack, *Scenario Planning in Organizations: How to Create, Use, and Assess Scenarios* (San Francisco, CA: Berrett-Koehler Publishers, Inc., 2011), pp. 10-13.

²⁴⁶ For examples see Peter Schwartz, *The Art of the Long View: Planning for the Future in an Uncertain World* (New York, NY: Doubleday, 1996); James A. Ogilvy, *Creating Better Futures: Scenario Planning as a Tool for a Better Tomorrow* (New York, NY: Oxford University Press, 2002); Kees Van Der Heijden, *Scenarios: The Art of Strategic Conversation* (New York, NY: John Wiley & Sons, Ltd., 2009); Mats Lindgren and Hans Bandhold, *Scenario Planning: The Link Between Future and Strategy* (New York, NY: Palgrave Macmillan, 2009); Alfred Marcus, *Strategic Foresight: A New Look at Scenarios* (New York, NY: Palgrave Macmillan, 2009); and Thomas J. Chermack, *Scenario Planning in Organizations: How to Create, Use, and Assess Scenarios* (San Francisco, CA: Berrett-Koehler Publishers, Inc., 2011).

²⁴⁷ Andrew F. Krepinevich, *7 Deadly Scenarios: A Military Futurist Explores War in the 21st Century* (New York, NY: Bantam, 2009), p. 10.

²⁴⁸ Joint Chiefs of Staff, *Joint Publication 5-0: Joint Planning Operations* (Washington, DC: Department of Defense, August 11, 2011), pp. III-36-37, and IV-27-36.

Scenario planning's success has been its ability to place decision makers and analysts in novel and unfamiliar circumstances, challenging the basis of their knowledge and expectations. Scenarios are not predictions, and generally not associated with distinctive probabilities regarding their likelihood of occurring. Instead, they constitute sets of qualitatively distinct future states of the world, each of which varies upon one or more dimensions.

Scenario developers have noted that trying to divine or predict a single outcome can be a disservice to consumers by reducing the number of alternative futures under consideration. Moreover, if analysts focus their attention on those cases that are estimated to be the most likely to occur, they may neglect less-likely cases with particularly important consequences. Analysts are helped by generating several scenarios because they focus attention on the key drivers and lynchpins that may produce different qualitative outcomes, expose critical assumptions, and reveal useful indicators for identifying when high impact/low probability outcomes are occurring.²⁴⁹

Scenarios have been used in multiple intelligence products, and included in teaching applications and outreach efforts.²⁵⁰ Because they are largely intellectual exercises designed to expose users to new ideas and situations, rather than advance particular judgments, they can facilitate exchanges of ideas among participants that might otherwise have difficulty sharing analytic products or data that include sensitive

²⁴⁹ Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), p. 123.

²⁵⁰ For examples see National Intelligence Council, *Global Trends 2025: A Transformed World* (Washington, DC: Directorate of National Intelligence, November 2008); and Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), pp. 124-131.

judgments or sources. This was demonstrated by the *Global Trends* series produced by the NIC, which employed scenarios to identify key drivers in the international system and provide policy makers and other readers with an understanding of the challenges ahead.²⁵¹ As Fingar noted, the production of scenario-based research can affect how the intelligence community can interact with outside experts and decision makers, enabling analysts to access alternative perspectives and insights that would be unavailable from more traditional intelligence sources.

When a few of us sat down in late 2007 to determine what we wanted to accomplish and how we would produce the fourth iteration of *Global Trends*, we recognized that we had an unprecedented combination of opportunity, experience, and willing participants. We knew that the 2008 election would result in an almost complete changeover of senior officials, no matter who won, and we saw this as a rare opportunity to help senior policy makers build their agendas with awareness of longer-term trends. In my experience, the start of a new administration is one of the few times officials have the time and appetite for strategic thinking; we were determined to hit that window of opportunity. Moreover, we had learned from our *Global Trends 2020* experience that non-Americans had much to contribute and would be influenced by our next look to the future. This offered an opportunity for focused dialog with influential people from many nations on issues that they and the Americans who would produce *Global Trends 2025* considered to be among the most important. This, in turn, offered opportunities for collaboration to address those and other issues. The 2025 report has been translated into Chinese, Korean, and probably other languages, and my former NIC colleagues and I have given dozens of public presentations in the United States and around the world.

We made one major procedural change in the way we prepared the 2025 report. This time, we invited both American and non-American contributors to comment on draft versions of the report. We solicited

²⁵¹ See National Intelligence Council, *Global Trends 2010*, http://www.dni.gov/nic/special_globaltrends2010.html (accessed on June 17, 2012); National Intelligence Council, *Global Trends 2015: A Dialogue About the Future With Nongovernment Experts* (Washington, DC: National Intelligence Council, December 2000); National Intelligence Council, *Mapping the Global Future: Report on the National Intelligence Council's 2020 Project Based on Consultations With Nongovernmental Experts Around the World* (Washington, DC: National Intelligence Council, December 2004); and National Intelligence Council, *Global Trends 2025: A Transformed World* (Washington, DC: National Intelligence Council, November 2008).

input through a variety of conferences, seminars, and commissioned studies, as we had done previously. But this time we also posted drafts on a special website and invited continuing interchange to ensure that we had understood points correctly, to smoke out alternative judgments, and to ensure that we were communicating judgments effectively. The process worked. We produced a better product, and we built interest in and support for the project among influential people around the world....

Global Trends 2025, like its predecessors, was designed to stimulate discussion, debate, and strategic thinking about the future.... Discussion of the issues we addressed by persons outside of the U.S. government is a good thing, but our primary audience was, of course, the incoming administration and career officials in Washington. I can write with confidence that the report was read and discussed and that it continues to serve as a frame of reference for discussion and decisions.²⁵²

In spite of its success, the *Global Trends* series revealed important weaknesses in scenario-based approaches. Because scenario planning seeks to place decision makers in challenging situations, scenario users and developers are often agnostic to, or unaware of, the causal mechanisms that determine how these situations were arrived at, thus disconnecting the preparation for different futures from strategies for shaping them as outcomes.²⁵³ *Global Trends 2025* envisioned many different drivers or trends that will shape the international system; however, these are not formally connected to different scenarios via a single causal, or collection of causal models. Although many drivers are specified, such as the increasing political, social, and economic mobilization of women and potential revolutions in medical technologies, their roles in driving the international system toward one and away from other examined scenarios are not.²⁵⁴

²⁵² Thomas Fingar, *Reducing Uncertainty: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), pp. 56-58.

²⁵³ Alfred Marcus, *Strategic Foresight: A New Look at Scenarios* (New York, NY: Palgrave Macmillan, 2009), pp. 9-10.

²⁵⁴ National Intelligence Council, *Global Trends 2025: A Transformed World* (Washington, DC: National Intelligence Council, November 2008), pp. 16-17, and 47-49.

Second, as the number of alternative futures expands, analysts and planners may be unable to generate enough cases to fully explore the range of the possible states of the system. Moreover, even if a large number of scenarios could be generated, few decision makers are able to work through the implications of more than a small set of alternatives. For example, a recent scenario-based study of the future security environment for the Department of Defense noted that approximately six scenarios were useful to decision makers before diminishing returns set in.

How many scenarios should be developed and analyzed? In theory, dozens could be written that would present significant security challenges for the United States, some more plausible than others. Yet senior defense officials cannot devote their attention to thirty or forty different contingencies, nor can they determine how to reshape the armed forces with its many possibilities in mind, especially if no effort is made to establish a clear priority among them. Perhaps more importantly, examining such a large number of scenarios in detail would quickly result in diminishing returns....

Scenarios are not intended to be predictions of the future, however. Rather, scenario-based planning is a method for minimizing risk. Scenario planners accomplish this by developing a number of plausible futures—perhaps half a dozen—that produce a comprehensive list of critical operational challenges. If senior leaders employ these scenarios to develop, assess, and create options for addressing these challenges, they can significantly minimize the prospect of being caught off guard or underprepared whenever and wherever new threats emerge.²⁵⁵

The dual problem of generating sufficient numbers of scenarios and evaluating them has been an enduring challenge since the scenarios were first employed. Indeed, reflecting upon his experiences with Royal Dutch Shell, Pierre Wack identified an even smaller number, four, that should be employed in decision-making processes.

²⁵⁵ Evan Braden Montgomery, *Defense Planning for the Long Haul: Scenarios, Operational Concepts, and the Future Security Environment* (Washington, DC: Center for Strategic and Budgetary Assessments, 2009), pp. 14-15.

Decision scenarios acknowledge uncertainty and aim at structuring and understanding it—but not by merely crisscrossing variables and producing dozens or hundreds of outcomes. Instead, they create a few alternative and internally consistent pathways into the future. They are not a group of quasi-forecasts, one of which may be right. Decision scenarios describe different worlds, not just different outcomes in the same world. Never more than four (or it becomes unmanageable for most decision makers), the ideal number is one plus two; that is, first the surprise-free view (showing explicitly why and where it is fragile) and then two other worlds or different ways of seeing the world that focus on the critical uncertainties.²⁵⁶

The four scenarios employed in *Global Trends 2025* were characterized by the study's lead authors as a means for challenging policy makers and other readers into thinking about possible futures, but that this list was a sample, not an exhaustive exploration of the possibilities.

In the four fictionalized scenarios, we have highlighted new challenges that could emerge as a result of the ongoing global transformation. They present new situations, dilemmas, or predicaments that represent departures from recent developments. As a set, they do not cover all possible futures. None of these is inevitable or even necessarily likely; but, as with many other uncertainties, the scenarios are potential game-changers.

- In **A World Without the West**, the new powers supplant the West as the leaders on the world stage.
- **October Surprise** illustrates the impact of inattention to global climate change; unexpected major impacts narrow the world's range of options.
- In **BRICs' Bust-Up**, disputes over vital resources emerge as a source of conflict between major powers—in this case two emerging heavyweights—India and China.
- In **Politics is Not Always Local**, nonstate networks emerge to set the international agenda on the environment, eclipsing governments.²⁵⁷

²⁵⁶ Pierre Wack, "Scenarios: Shooting the Rapids," *Harvard Business Review*, No. 85617 (November-December 1985), p. 9.

²⁵⁷ National Intelligence Council, *Global Trends 2025: A Transformed World* (Washington, DC: National Intelligence Council, November 2008), pp. xii-xiii.

Given the large number of drivers, the number of potential scenarios that policy makers should be concerned with was astronomically higher, leaving intelligence producers and consumers with questions as to whether the four examined scenarios were the most interesting, stressful, or revealing given the unknown features of other cases that were not considered. For example, in the area of emerging technologies alone, nine technologies and the drivers and barriers to their development and diffusion were considered:

- Ubiquitous computing;
- Clean water;
- Energy storage;
- Biogerontechnology;
- Clean coal;
- Human strength augmentation;
- Biofuels;
- Service robotics;
- Human cognitive augmentation.²⁵⁸

A simple understanding of combinatorics reveals just how small of the possibility space four scenarios actually are. If each of these technologies was coded as a simple binary variable, e.g. mature or immature, there would be more than five hundred distinct scenarios—meaning that these four scenarios represent less than one percent of possible outcomes. While many of these scenarios may be duplicative or uninteresting, how to determine this *a priori* is unclear.

Although scenarios are intended to assist analysts and decision makers cope with uncertainty through generating and experiencing alternate worlds, they nevertheless succumb to the same cognitive challenges as other techniques. Restricting the number of scenarios to a manageable, small number limits the extent to which the dimensions of

²⁵⁸ National Intelligence Council, *Global Trends 2025: A Transformed World* (Washington, DC: National Intelligence Council, November 2008), pp. 47-49.

complex systems can be explored, and compels analysts to choose what to include and exclude from consideration based on their *a priori* assumptions of importance. As a result, the pruning of potential scenarios or combinations of variables occurs prior to scenario development, and depends on the preexisting mental models of scenario developers in order to determine what warrants attention and what does not. The step-wise search across alternative drivers that characterize different scenarios may miss cases where the most interesting or challenging cases depend on complex combinations of factors that may be missed by manually considering each variable in isolation.

Scenario-based planning does grapple with the multiplicity of plausible futures but has two important weaknesses for LTPA [Long Term Policy Analysis]. First, the choice of any small number of scenarios to span a highly complex future is ultimately arbitrary. A scenario exercise will necessarily miss many important futures that do not make the cut into the top few. Despite best efforts, the logic used to sort the scenarios may seriously bias any conclusions drawn from them. As one small example, research in the psychology of decisionmaking indicates that humans gravitate to stories whose plot revolves around a single dramatic event rather than to those where the ending is driven by the slow accumulation of incremental change.... Thus, scenario-based planning exercises may make it difficult to think about responses to slowly emerging problems. Certainly they will fail to address many of the challenges that the future will hold.²⁵⁹

Although scenarios may be helpful for exploring discontinuities resulting from ongoing trends or potential developments, they suffer from the same challenges as other analytic methods, such as case study research: As the number of variables under consideration increase, the size of the potential space of possible cases grows exponentially, outstripping the number of historical cases available for study.²⁶⁰

²⁵⁹ Robert J. Lempert, Steven W. Popper, and Steven C. Bankes, *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis* (Santa Monica, CA: RAND, 2003), p. 35.

²⁶⁰ Sean J. Edwards, *Swarming and the Future of Warfare* (Santa Monica, CA: RAND, 2004), p. 22.

Given these difficulties, conventional scenario-planning techniques can be augmented by the use of formal mathematical or computational models. These models can generate massive quantities of cases, including those derived from alternative theories, sets of initial parameters, and complex combinations of variables and interactions. While these model results lack the narrative depth of traditional, manually generated scenarios, their ability to generate hundreds, thousands, or even millions of outcomes can ensure that analysts perform a more thorough search across alternative futures. This then enables the most interesting or challenging cases to be selected for further consideration and serve as a template for the development of more traditional narrative scenario-based analysis.²⁶¹

Alternative Competing Hypotheses

ACH has developed into a staple of analytic tradecraft. Whereas scenarios look toward the future, extending beyond the evidence at hand to explore future possibilities, ACH is a technique for examining information and determining its diagnostic value in supporting or refuting alternative hypotheses about an intelligence target's behavior or future. While its development resulted from the initial work of the Methods and Forecasting Division during the 1970s, similar approaches were also recommended by

²⁶¹ Robert J. Lempert, David G. Groves, Steven W. Popper, and Steve C. Bankes, "A General, Analytic Method for Generating Robust Strategies and Narrative Scenarios," *Management Science*, Vol. 52, No. 4 (April 2006), pp. 514-528; David G. Groves and Robert J. Lempert, "A New Analytic Method for Finding Policy-Relevant Scenarios" *Global Environmental Change*, Vol.17, No. 1 (2007), pp. 73-85; and Robert Lempert, "Can Scenarios Help Policymakers Be Both Bold and Careful?" in Frances Fukuyama, ed., *Blindside: How to Anticipate Forcing Events and Wild Cards in Global Politics* (Washington, DC: Brookings, 2007), pp. 109-119.

other scholars and practitioners in other countries, e.g. Ben-Israel's amended critical method to employ the scientific method in intelligence estimates.²⁶²

ACH formalized an essential element of analytic tradecraft by linking evidence and inference under uncertainty, and provided a firm grounding for analytic practice in cognitive science and the philosophy of science in cases where events are either underdetermined or overdetermined, thus allowing for many alternative explanations or expectations about current and future events. As already noted, ACH was specifically developed to expand the search for alternative explanations of events and minimize the impact of bounded rationality regarding analytic satisficing in the generation and testing of alternative hypotheses.

From a scientific perspective, ACH resembles the scientific practice of Inference to the Best Explanation (IBE) where scientists evaluate evidence for which many explanations may exist. In these cases, IBE provides a basis for choosing between alternative competing explanations for an observed outcome by selecting the simplest or most likely explanation.²⁶³ Importantly, IBE does seek to make explanatory inferences for particular data, rather than devise a theory applicable to a general class of cases, and is therefore consistent with the needs of policy makers and analysts who work on specific problems rather than general classes of them.

Compared with IBE, however, ACH is distinct. IBE seeks to eliminate weaker explanations from stronger ones, and thus depends heavily on the preexisting beliefs or

²⁶² Isaac Ben-Israel, "Philosophy and Methodology of Intelligence: The Logic of Estimate Process," *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 708-714.

²⁶³ Samir Okasha, *Philosophy of Science: A Very Short Introduction* (New York, NY: Oxford University Press, 2002), pp. 29-30.

experiences of the scientist employing it to provide the selection criterion. Often, scientists prefer selecting the simplest explanation, such as in physical systems.²⁶⁴

Alternatively, in systems that have evolved over time, such as biological and social systems, there is no reason to believe that the simplest explanation is the most likely.²⁶⁵

In physics, it is a fair principle that the simplest model for any particular phenomenon is usually the right one. But in biology, accidents of history often invalidate this principle. It is only the improbability of very complicated arrangements that have been reached by biological evolution that makes a criterion of simplicity at all relevant. And in fact it may no more be possible to understand the construction of a biological organism than a computer program: each is arranged to work, but a multitude of arbitrary choices is made in its construction.²⁶⁶

In the case of the intelligence community, where the challenges analysts face often involve understanding complex social, financial, military, and leadership structures and processes, selecting the best explanation based on simplicity may lead to highly misleading conclusions. For example, it may be easy to conclude that a foreign leader is irrational, self-destructive, and temperamental, rather than discern a complex web of historical experiences and secret agreements which, when violated, provoked anger and defiance.²⁶⁷ Indeed, intelligence agencies, military organizations, and other strategic actors operate according to a paradox that limits and eschews direct, simple, and straightforward action.²⁶⁸

²⁶⁴ Hilary Putnam, *The Collapse of the Fact/Value Dichotomy and Other Essays* (Cambridge, MA: Harvard University Press, 2002), pp. 30-43.

²⁶⁵ John Lewis Gaddis, *The Landscape of History: How Historians Map the Past* (New York, NY: Oxford University Press, 2002), pp. 57-66.

²⁶⁶ Stephen Wolfram, *Cellular Automata and Complexity* (Boulder, CO: Westview Press, 1994), pp. 493-494.

²⁶⁷ This point was made in a discussion with H. Stapleton Roy, Woodrow Wilson Institute, June 12, 2011.

²⁶⁸ Edward N. Luttwak, *Strategy: The Logic of War and Peace* (Cambridge, MA: The Belknap Press of Harvard University Press, 2001), pp. 3-15.

The problem for IBE is that if no objective or external criteria exists for choosing between alternative explanations, then the choice of what constitutes the “best” explanation becomes arbitrary. Moreover, determining what constitutes the best explanation is contingent upon competing preferences for general causes that apply to entire classes of phenomena or contingent ones that apply to particular cases.²⁶⁹ As Godfrey-Smith noted, many scientists employ IBE in the search for the single, best explanation, rather than discover the entire universe of possible explanations, an approach he believed is misguided.

The term “explanatory inference” suggests that there is one kind of relationship between data and hypothesis—the explanation relationship—that is involved in explanatory inference. Many philosophers would accept this. The term “inference to the best explanation” is, in fact, a more common name for what I call explanatory inference; that term suggests that there is some single measure of “explanatory goodness” involved. But I think this is the wrong way to think about scientific reasoning (and this is why I have avoided the term “inference to the best explanation”). I use “explanatory inference” in a broader way that does not suppose that there is a single measure of explanatory goodness involved, which applies to all of science. Rather, explanatory inference is a matter of devising and comparing hypotheses about hidden structures that might be responsible for data. “Explanation” is seen as something pretty diverse.²⁷⁰

Thus, IBE is not immune to the cognitive biases that persistently challenge analysts, such as mirror-imaging or substitution noted earlier. This was demonstrated in an example provided by Samir Okasha regarding whether a mouse or a maid is most likely responsible for missing cheese after hearing scratching sounds. Okasha questioned why

²⁶⁹ Nancy Cartwright, *How the Laws of Physics Lie* (New York, NY: Oxford University Press, 2002), pp. 15, and 87-99.

²⁷⁰ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), p. 199.

the mouse is assumed to be the better explanation, and concluded that it remained embedded in the premises or assumptions of the observer.

Why do we regard the mouse hypothesis as a better explanation of the data than the maid hypothesis? Presumably, because we know that maids do not normally steal cheese, whereas mice do. But this is knowledge that we have gained through ordinary inductive reasoning, based on our previous observations of the behaviour of mice and maids. So according to this view, when we try to decide which of a group of competing hypotheses provides the best explanation of our data, we invariably appeal to knowledge that has been gained through ordinary induction. Thus it is incorrect to regard IBE as a more fundamental mode of inference. ...

If we want to use IBE, we need some way of deciding which of the competing hypotheses provides the best explanation of the data. But what criteria determine this? A popular answer is that the best explanation is the simplest or the most parsimonious one....

...But if scientists use simplicity as a guide to inference, this raises a problem. For how do we know that the universe is simple rather than complex? Preferring a theory that explains the data in terms of the fewest number of causes does seem sensible. But is there any objective reason for thinking that such a theory is more likely to be true than a less simple theory?²⁷¹

In contrast to IBE, ACH is focused on preserving the lifespan of alternative hypotheses by keeping those that cannot be eliminated by the available information in the minds of analysts. Thus, rather than select the best hypothesis among alternatives based on the strength of the evidence in its favor, ACH is focused on eliminating those that cannot withstand scrutiny.

ACH involves identifying a set of mutually exclusive alternative explanations or outcomes (presented as hypotheses), assessing the consistency or inconsistency of each item of evidence with each hypothesis, and selecting the hypothesis that best fits the evidence. The idea behind the technique is to refute rather than to confirm each of the hypotheses. The most likely hypothesis is the one with the least evidence

²⁷¹ Samir Okasha, *Philosophy of Science: A Very Short Introduction* (New York, NY: Oxford University Press, 2002), pp. 29, and 32-33.

against it, as well as evidence for it, not the one with the most evidence for it.²⁷²

An additional strength of ACH is its ease of use. When originally articulated, ACH consisted of a table where evidence was arranged on the left-hand column, followed by additional columns stating each alternative hypothesis to the right. In each cell, evidence was evaluated as highly consistent, consistent, not applicable, inconsistent, or highly inconsistent with the hypotheses in question. Working across the rows evaluated the overall diagnosticity of each data element, i.e. whether or not it aids analysts in determining the relative merits of alternative hypotheses. Once completed, evidence that had no diagnostic value, i.e. information that is consistent with all alternatives, can be removed from consideration. Working down the columns assesses the relative merits of each hypotheses by allowing for an assessment of the weight against it. This arrangement is shown in Table 3-1 below:

²⁷² Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), p. 160.

Table 3-1: Original Alternative Competing Hypotheses Table.

Evidence	H1	Hn
Data Element 1	++	++
Data Element 2	+	+
Data Element 3	NA	NA
Data Element 4	-	-
Data Element 5	--	--
Data Element n

Note. Table adapted from Richards J. Heuer, Jr., *The Psychology of Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 1999), p. 101.

Importantly, the emphasis on evaluating hypotheses is on their refutation, consistent with Karl Popper’s famous demarcation criterion and in recognition of the confirmation bias noted earlier in this chapter.

In evaluating the relative likelihood of alternative hypotheses, start by looking for evidence or logical deductions that enable you to reject hypotheses, or at least to determine that they are unlikely. A fundamental precept of the scientific method is to proceed by rejecting or eliminating hypotheses, while tentatively accepting only those hypotheses that cannot be refuted. The scientific method obviously cannot be applied *in toto* to intuitive judgment, but the principle of seeking to disprove hypotheses, rather than confirm them, is useful.

No matter how much information is consistent with a given hypothesis, one cannot prove that hypothesis is true, because the same information may also be consistent with one or more other hypotheses. On the other hand, a single item of evidence that is inconsistent with a hypothesis may be sufficient grounds for rejecting that hypothesis. . . .

In examining the matrix, look at the minuses, or whatever other notation you used to indicate evidence that may be inconsistent with a hypothesis. The hypothesis with the fewest minuses is probably the most likely one. The hypothesis with the most minuses is probably the least

likely one. The fact that a hypothesis is inconsistent with the evidence is certainly a sound basis for rejecting it. The pluses, indicating evidence that is consistent with a hypothesis, are far less significant. It does not follow that the hypothesis with the most pluses is the most likely one, because a long list of evidence that is consistent with almost any reasonable hypothesis can be easily made. What is difficult to find, and is most significant when found, is hard evidence that is clearly inconsistent with a reasonable hypothesis.²⁷³

More recent versions of ACH have added additional information to the approach, standardizing language for determining the diagnosticity of data elements, and adding metadata regarding each data element, such as its source and estimated credibility. Moreover, software tools have been developed to aid analysts with breaking complex hypotheses into sets of mutually exclusive conjectures, identifying data that lacks diagnosticity, sorting information or hypotheses based on user specified criteria, and allowing multiple users to collaborate on developing the content of ACH tables and comparing their independent assessments in order to identify where analysts agree or disagree.²⁷⁴ A contemporary ACH table showing the range of values that may be located in each column is below in Table 3-2:

²⁷³ Richards J. Heuer, Jr., *The Psychology of Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 1999), p. 104.

²⁷⁴ A computerized version of ACH was demonstrated to the author as part of the ThinkSuite software toolkit by Randolph Pherson and Mary Boardman of Pherson Associates, LLC on April 26, 2012. Additional information about these tools is available at <http://www.pherson.org/Tools.html> (accessed on June 18, 2012). Also see Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly, 2011), pp. 164-166.

Table 3-2: Contemporary Alternative Competing Hypotheses Table.

Evidence	Source Type	Credibility	Relevance	H1	Hn
Data Element 1	Assumption	High	High	Compelling	Compelling
Data Element 2	HUMINT	Medium	Medium	Consistent	Consistent
Data Element 3	SIGINT	Low	Low	NA	NA
Data Element 4	IMINT	Inconsistent	Inconsistent
Data Element 5	MASINT	Refuting	Refuting
Data Element 6	GEOINT
Data Element 7	OSINT
Data Element 8	Liaison Reporting
Data Element 9	Lack of Intelligence Reporting Despite Vigorous Search
Data Element 10	All-Source Analysis				
Data Element 11	Inference				
Data Element n	Contrarian Hypothesis				

Note. Table adapted from Richards J. Heuer, “Computer-Aided Analysis of Competing Hypotheses,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 254-255.

Important additions to ACH, as demonstrated above, include the addition of assumptions; judgments of other analytic assessments such as the potential presence of adversarial denial or deception, e.g. Iraqi WMD; and contrarian hypotheses.

Interestingly, contrarian hypotheses may be included as either an alternative hypothesis,

meriting its own column, or as evidence that affects how other hypotheses should be evaluated. For example, the hypothesis that Iraq's WMD programs were in fact suspended prior to the 2003 Iraq War, but were intended to appear active in order to deter the regime's domestic and regional rivals could be treated as a distinct explanation worthy of representation—as a column against which other data is evaluated, or as a data element to be included as a row if the assumed evidence was the existence of a denial and deception program.²⁷⁵

When viewed collectively, the evolution of ACH starting from 1999 and its contemporary form has been the inclusion of assumptions, missing information, denial and deception, and prior judgments into the matrix in order to ensure they are considered. These additions have extended ACH from an exclusively empirical technique designed to assess the diagnosticity of available information, into a structured framework for evaluating collection, counterintelligence, and analysts' imagination in addition to empirical evidence.

ACH has many strengths, but also important limitations. While ACH has expanded to operate beyond the limitations of empiricism, it remains primarily data driven. As a result, its application to problems in which analysts have little or no information is limited. Instead, ACH is tailored for problems where information is ambiguous and incomplete, but can run into difficulty when information is simply missing. Unfortunately, this situation is more common in intelligence than producers

²⁷⁵ Richards J. Heuer, "Computer-Aided Analysis of Competing Hypotheses," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 254-255.

would like, because consumers drive priorities and rarely wait for producers to develop the necessary collection capabilities and analytic expertise before making policy decisions.

Just as analysts cannot very well ignore an issue that policymakers want analyzed, they also cannot say, “We have no collection on that, so we’re not analyzing it.” Quite simply, analysts are expected to analyze despite the paucity of collection. Indeed, since it is widely recognized that there will never be perfect or complete collection on any issue, there is always going to be some part of analysis that is based more on analytical experience and judgment than on solid information.²⁷⁶

In such cases, analysts may simply lack the evidentiary basis against which ACH can be effectively utilized.

A second limitation of ACH is that it remains dependent on the imagination of analysts. Thus, it is only capable of evaluating the alternative hypotheses that analysts articulate and consider. When confronting highly complex intelligence challenges, the number and diversity of alternative explanations will almost always exceed the limits of what individual or groups of analysts can imagine. As Thomas Schelling famously noted, “One thing a person cannot do, no matter how rigorous his analysis or heroic his imagination, is to draw up a list of things that would never occur to him.”²⁷⁷ Likewise, Ben-Israel noted that while science was fundamentally about the testing of hypotheses, their generation was outside of its bounds, belonging to a different domain of knowledge, expertise, and creativity.

²⁷⁶ Mark M. Lowenthal, “Intelligence Analysis: Management and Transformation Issues,” in Jennifer E. Simms and Burton Gerber, eds., *Transforming U.S. Intelligence* (Washington, DC: Georgetown University Press, 2005), p. 232.

²⁷⁷ Quoted in Mark Herman, Mark Frost, and Robert Kurz, *Wargaming for Leaders: Strategic Decision Making from the Battlefield to the Boardroom* (New York, NY: McGraw Hill, 2009), p. 3.

The starting-point of the scientist is not, as the layman believes, “cool” and impartial derivation of theories, laws and predictions from precise facts; rather, theories are, as Popper says: “*free* creations of our minds, the result of an almost poetic intuition”.... Or in the words of Albert Einstein: “There is no logical path leading to these... laws. They can only be reached by intuition, based upon something like an intellectual love of the objects of experience.” What does this imply for the desired character of the estimator?

One thing is clear: the estimator must have a *creative imagination*. Otherwise, he will fail right at the beginning—in the starting phase of creating possible hypotheses.²⁷⁸

Indeed, Heuer’s assessment of whether ACH might have affected the 2002 NIE of Iraqi WMD capabilities is particularly revealing. Specifically, he noted that the available historical experience did not suggest that Saddam Hussein’s weapons programs had been dismantled and only persisted as a deception campaign. Instead, the set of alternatives focused on competing assessments of their size, scope, and technical features. Hypotheses that suggested that these programs did not exist and were part of a deception campaign were simply not in the set that analysts considered. Simply by including the possibility that Iraq’s WMD programs had been dismantled the entire structure of an ACH-based analysis would have changed by revealing the resilience of this hypothesis, never supported by the available evidence but never refuted either.

Another weakness of ACH is the problem of complexity. The diagnostic emphasis of ACH assumes that analysts know when data is consistent or inconsistent with particular hypotheses. However, the study of complex adaptive systems has shown there are many classes of problems in which the cognitive limitations of the human brain prevent accurate assessments and lead to flawed expectations. When problems contain

²⁷⁸ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 713.

sufficient complexity, analysts may easily mischaracterize relationships between evidence and inference, data and hypothesis. If this occurs, information consistent with a hypothesis may be attributed as inconsistent, while information that is inconsistent is mistaken to be consistent.

The issue of hypothesis generation and evaluation, then, constitutes important limitations to ACH, and an area where computational modeling can provide significant assistance to analysts. Although computation is normally associated with the analysis of data, many machine learning and evolutionary computation techniques originally developed by the Artificial Intelligence (AI) community have demonstrated how algorithms can generate and test more models and hypotheses than researchers can do manually. By automating the generation and testing of models, computation can recommend candidate hypotheses for analysts to consider.²⁷⁹ In fact, because of its ability to generate unexpected scenarios, ABM has been regarded as a particularly effective tool for augmenting producers and consumers' imagination, aiding them in thinking about new hypotheses and relating them to empirical evidence of synthetic observations of artificial societies.²⁸⁰

The Challenge of Strategic Intelligence

At the start of this chapter, intelligence failures were identified as resulting from two sources: mindsets and an imbalance between strategic intelligence vs. current intelligence. While the previous sections examined mindsets and efforts to cope with

²⁷⁹ Pat Langley, Herbert A. Simon, Gary L. Bradshaw, and Jan M. Zytkow, *Scientific Discovery: Computational Explorations of the Creative Process* (Cambridge, MA: MIT Press, 1987); and Herbert A. Simon, *The Sciences of the Artificial* (Cambridge, MA: MIT Press, 1996).

²⁸⁰ Interview with Leon Fuerth, National Defense University, September 21, 2011.

them, most notably the development of SATs in recognition that analysis is a type of modeling activity that copes with cognitive biases, the problem of strategic intelligence is an institutional one.

Strategic intelligence, sometimes referred to as “strategic analysis,” is generally regarded as Kent’s estimative intelligence and other analytic products that look into the future, provide context for framing decisions or interpreting current reporting, and speculate about the causal relations in the system that may reveal the benefits or dangers of particular activities or opportunities. Strategic intelligence primarily deals with long-term, enduring issues that policy makers must cope with, including efforts to resolve deep uncertainty surrounding intelligence problems and in cases where additional collection is not feasible or capable of providing information of diagnostic value. Whereas current intelligence provides consumers with descriptions of what is happening in the world, strategic intelligence seeks to raise their awareness of possible scenarios and the potential capabilities and intentions of others.

Lowenthal provided an illuminating distinction between tactical and strategic surprise, capturing the essential distinction between current and strategic intelligence, and characterized the differences between describing events in the world and recognizing the possibility of their occurrence.

Putting the difference between the two types of surprise in perspective, suppose, for example, that Mr. Smith and Mr. Jones are business partners. Every Friday, while Mr. Smith is lunching with a client, Mr. Jones helps himself to money from the petty cash. One afternoon Mr. Smith comes back from lunch earlier than expected, catching Mr. Jones red-handed. “I’m surprised!” they exclaim simultaneously. Mr. Jones’s surprise is tactical: He knew what he was doing but did not expect to get caught. Mr.

Smith's surprise is strategic: He had no idea the embezzlement was happening.²⁸¹

Current intelligence provides consumers with descriptions of what is occurring at any given moment; strategic intelligence, however, is more speculative and characterizes possibilities so that consumers can prepare themselves for events that might happen. Therefore, strategic intelligence provides a context or framework for understanding what is possible, while current intelligence describes what is happening. From this perspective, strategic intelligence is not necessarily a matter of providing intelligence on "important things" but making consumers aware of the "kinds of things" that are possible, what their importance might be if they occur, and how they could come to pass.

Another way that the intelligence community thinks about strategic intelligence is as mysteries that must be interpreted and divined, rather than puzzles that are solved via the collection of their pieces. This, again, emphasizes the epistemological limitations of a purely empirical intelligence analysis, as increased collection can assist with but not solve mysteries.²⁸² Often, these limitations affect consumers' expectations of intelligence analysis, particularly when their characterization of an intelligence problem differs from those of producers or they are simply unaware of how intelligence collection and analysis operate.²⁸³

Policy makers do not always appreciate the limits of what can be collected and known with certainty, the reasons behind ambiguity, and, occasionally, the propriety of intelligence. They sometimes confuse the lack of a firm estimate with pusillanimity when that may not be the case.

²⁸¹ Mark M. Lowenthal, *Intelligence: From Secrets to Policy* (Washington, DC: Congressional Quarterly Press, 2010), p. 3.

²⁸² Loch K. Johnson, "Introduction," in Loch K. Johnson, ed., *Handbook of Intelligence Studies* (New York, NY: Taylor and Francis, 2009), pp. 2-3.

²⁸³ Interview with Leon Fuerth, National Defense University, September 21, 2011.

Intelligence officers sometimes liken this problem to the difference between puzzles and mysteries. Puzzles have solutions; these may be difficult but they can be found. Mysteries, on the other hand, may not have a knowable solution. This distinction may be lost on policy makers but it is very real in the minds of intelligence officers. They expect to be asked to solve puzzles; they know they may not be able to solve mysteries.²⁸⁴

The estimative or speculative character of strategic intelligence is ultimately produced via a combination of experience, expertise, and tradecraft. Because strategic intelligence estimates rely on inference rather than evidence, strategic intelligence will always be vulnerable to dismissal from consumers. As a result, consumers have far less of an appetite for strategic intelligence than they do for current intelligence. Hilsman, Kent, and others noted that strategic intelligence often marks the low point of consumer interest in intelligence products, and generally antagonizes producer/consumer relations. Thus, despite repeated calls for more strategic intelligence as a way of improving the overall quality of analysis, as a practical matter, consumers have little demand for it. As Fingar explained, intelligence producers confront a dilemma where they cannot earn the trust and confidence of consumers if they do not satisfy their demands for current intelligence, but they cannot adequately provide strategic warning about threats and opportunities unless they explore issues that sit outside of the immediate needs and interests of consumers.

Commentary on the Intelligence Community often draws a distinction between “current” and “strategic” intelligence, usually to decry excessive attention to explaining the latest intelligence factoids obtained by collectors and inadequate attention to longer-term strategic analysis. Such criticisms are valid, but they generally miss (or misrepresent) important

²⁸⁴ Mark M. Lowenthal, *Intelligence: From Secrets to Policy* (Washington, DC: Congressional Quarterly, 2010), p. 193.

points. One error is to underestimate the demand for what the military calls “situational awareness.”

Policy makers throughout the government want to know what is happening in their geographic and/or substantive areas of responsibility.... To prevent surprise and embarrassment, policy makers expect their own staff and their Intelligence Community support team to ensure that they are always “on top of” their own portfolios. This is the demand-side “pull” for current intelligence. There is also a supply-side “push” from the IC support team. Knowing that their ability to provide timely, targeted, and useful intelligence support to their primary customers requires winning and maintaining the customer’s confidence, IC analysts err on the side of providing more “current intelligence” than necessary. Some policy makers will say, from time to time, that they would welcome more long-range analysis, but the unspoken caveat is that receiving it should not come at the expense of constant situational awareness.

Even though policy makers often are so consumed by the demands of their in-box and immediate policy issues that they have little time or appetite for longer-term developments, the pundits who proclaim the need for more “strategic” analysis are right. In the grand scheme of things, it is far more important for the Intelligence Community to reduce uncertainty about what might happen in the future than it is to ensure that policy makers know what has happened in the last few days or hours. Indeed, it can be argued that the most important justification for having an intelligence enterprise is to provide “strategic warning” sufficiently far in advance that policy makers can act to prevent, ameliorate, or capitalize on the anticipated developments.²⁸⁵

Strategic intelligence constitutes the basic research and intellectual preparation that allows for analysts to adapt to rapidly changing circumstances. By investigating alternative futures, imagining scenarios, and considering as many possible perspectives on a given situation, analysts can prepare themselves to make sense of new information and unforeseen events, limiting the opportunities for strategic surprise while enabling

²⁸⁵ Thomas Fingar, *Reducing Intelligence: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), pp. 52-53.

rapid recovery from tactical ones. Thus, strategic intelligence provides the bedrock of analysts' expertise, training, and experience.²⁸⁶

The relationship between strategic and current intelligence constitutes one of the most problematic challenges in the management of intelligence. Precisely how intelligence resources—collection assets and analysts' attention—should be allocated is a complex and perpetual challenge whenever resources are limited. Even in cases of relative strategic clarity, such as the presumably simpler periods of World War II and the Cold War, the problems of resource allocation within the intelligence community were acute and enduring. For example, in 1949, shortly after the conclusion of World War II and the opening of the Cold War, Kent wrote:

Intelligence must be equipped to deal with the array of subjects which I will consider, and in the course of the years it may conceivably deal with all of the subjects at least once. It will, however, tend to with any single subject only when that subject is part of a threat to our national interest or is required by a prospective course of action. One of the most continuously vexing problems in the administration of intelligence is deciding which particular subjects shall be watched, reported upon, or made the object of descriptive or speculative research. Equally vexing is deciding the order of their priority. The point is that intelligence is always fully occupied, but occupied almost exclusively on a relatively few subjects of real national concern. At the same time intelligence must be ready to handle a large number of subjects.²⁸⁷

The persistent challenges of managing the tradeoff between current and strategic intelligence have never subsided, and have resided at the core of several intelligence failures in which analysts and policy makers were surprised by the existence of particular groups, conduct of dangerous activities, and by possibility that a given outcome could

²⁸⁶ John Hollister Hedley, "Analysis for Strategic Intelligence," in Loch K. Johnson, *The Oxford Handbook of Intelligence* (New York, NY: Taylor and Francis, 2009), p. 211.

²⁸⁷ Sherman Kent, *Strategic Intelligence for American World Policy* (Princeton, NJ: Princeton University Press, 1949), pp. 8-9.

occur. As Lowenthal suggested, these intelligence failures are distinctly different from failures of current intelligence, in which collectors and analysts search for specific information but are unable to gather or infer what they need. Pillar noted this problem when he argued that the intelligence community's failure to identify the Aum Shrinriko group that attempted to manufacture and disperse sarin gas on the Tokyo subway system in 1995 constituted a failure of strategic, not current, intelligence.

At the time of the attack, the group was not on the screens of the U.S. counterterrorist community. That fact represents a strategic intelligence failure—quite the opposite of the attention the intelligence community already was giving, even at that early date, to bin Ladin.... The failure did not receive attention because it did not involve the deaths of Americans or a large number of deaths of anyone else. But it might have. Aum Shinrikyo's bizarre ideology envisioned an Armageddon between Japan and the United States, which the group thought it might have to stimulate. If the Japanese police had not begun disrupting Aum's activities (the proximate stimulus for the attack on the subway), there is no telling how much lethal anti-U.S. mayhem the group might have caused.²⁸⁸

The scarce resources devoted to the problem of terrorism in 1995 had been mostly devoted to OBL and AQ, leaving little remaining to search for additional groups and alternative sources of religiously motivated terrorism. Thus, the surprise was not a failure to collect intelligence on other groups, it was the failure to consider their existence at all.

Balances between current and strategic intelligence also affect how intelligence agencies manage their staff and collection operations because they demand long-term, focused investments in personnel, technology, and operations before their value can be known to consumers. Thus, an extensive number of managerial decisions within the

²⁸⁸ Paul R. Pillar, *Intelligence and U.S. Foreign Policy: Iraq, 9/11, and Misguided Reform* (New York, NY: Columbia University Press, 2011), pp. 301-302.

intelligence community must be made in the absence of policy maker guidance regarding their anticipated needs and interests. Decisions regarding the hiring of individuals with particular expertise into intelligence organizations and developing necessary collection capabilities to be employed against future targets must be made years before the need for their capabilities manifests. The search for a balance between strategic and current intelligence undermines the dominant model of intelligence processes, the intelligence cycle, precisely because it challenges the assumption that policy makers have the foresight to shape intelligence collection and analysis years ahead in order to enable tailored capabilities to be developed in advance of their projected needs.²⁸⁹

Understanding the tradeoffs between current and strategic intelligence replicates the very cognitive difficulties associated with bounded rationality, and satisficing in the development and exploration of mental models addressed by SATs. Meeting the current intelligence needs of consumers denies analysts and their managers the opportunities to reflect on past performance—successes and failures—in order to improve their tradecraft and assess how resources have been allocated. By overcommitting current intelligence, intelligence analysts may lose the invaluable feedback from self-assessment as well as the consideration of alternative issues that have not yet captured the interest of policy makers. As George and Bruce noted:

[M]anagement imperatives that are driven by *current intelligence* demands (as opposed to more in-depth research and less time-pressured analysis)...do not permit sufficient time to reflect on the intelligence

²⁸⁹ Arthur S. Hulnick, “The Intelligence Cycle,” in Loch K. Johnson and James Wirtz, eds., *Intelligence: The Secret World of Spies* (New York, NY: Oxford University Press, 2011), pp. 65-66.

community's past performance or to record the lessons learned, from which subsequent generations of analysts can benefit.²⁹⁰

Likewise, Richard Russell noted the importance of management and organizational design in the development of strategic intelligence, and suggested that many of the reforms over the last decade missed an important opportunity to create new strategic intelligence organizations that did not need to respond to the daily, current intelligence demands of consumers.

A significant omission in the DNI's reforms is the creation of an IC strategic studies center. The Presidential Commission on Weapons Proliferation recommended the creation of "at least one not-for-profit 'sponsored research institute' to serve as a critical window into outside expertise for the Intelligence Community" and that its "primary purpose would be to focus on strategic issues." Such a center would be separated from the taxing burden of current intelligence production, have more expertise, and be better positioned than the CIA rank-and-file to fuse the information flows from public and clandestine sources to form strategic intelligence assessments.²⁹¹

Given that analysts have limited attention, and intelligence organizations have limited resources, the more intelligence organizations strive to meet the current intelligence needs of consumers, the less they can devote to strategic intelligence. From the perspective of evolutionary theory, this tradeoff is a version of exploration vs. exploitation.

The notion of exploration and exploitation has often appeared in intelligence studies with respect to the tradeoff between intelligence collection and policy actions, e.g. military operations or diplomatic activities that exploit information gained from

²⁹⁰ James B. Bruce and Roger Z. George, "Introduction: Intelligence Analysis—The Emergence of a Discipline," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 5.

²⁹¹ Richard L. Russell, "Achieving All-Source Fusing in the Intelligence Community," in Loch K. Johnson, ed., *Handbook of Intelligence Studies* (New York, NY: Taylor and Francis, 2009), p. 197.

intelligence sources and methods. In this context, tradeoffs exist between taking action against a target based on what is currently known vs. continuing to observe its behavior in order to gather new information, leading to potentially new and important discoveries.

For example, law enforcement or military operations may capture or kill the members of a single terrorist cell, while tracking and monitoring the cell's members may reveal more information about the larger group's intent, capabilities, organizational structure and operational decision making. Too great an emphasis on exploitation may prevent individual attacks, but may be akin to the game of "whack-a-mole" whenever operatives who have been captured and killed can be easily replaced while more senior members and operational capabilities are never discovered or affected. Alternatively, too great an emphasis on exploration may fail to prevent attacks despite sufficient warning, while increasingly extensive intelligence-gathering efforts may be discovered and denied via Operations Security (OPSEC) or even coopted into becoming sources of disinformation. These difficult and familiar tensions were evident during the decision making and planning of the raid on OBL's compound in 2011.²⁹²

Within intelligence analysis, framing of the problem of strategic vs. current intelligence views the time and attention devoted to developing and exploring mental models as the subject of tradeoffs between exploration and exploitation.²⁹³ Strategic intelligence is then understood to be a form of exploration that identifies new issues and topics that may eventually interest consumers, while current intelligence exploits what

²⁹² Peter L. Bergen, *Manhunt: The Ten-Year Search for Bin Laden: From 9/11 to Abbottabad* (New York, NY: Crown, 2012), pp. 122-135 and 163-195.

²⁹³ The issue of managing the attention of analysts and its relationship between current and strategic intelligence was raised by Paul Pillar during an interview conducted as part of this research project. Interview with Paul Pillar, February 2, 2012.

producers know about the interests of policy makers. As in other evolutionary systems, if a balance is not reached between the two, collapses may occur, thus viewing strategic surprise as a type of extinction event in biological systems.

Too much devotion to strategic intelligence results in a constant search for new ideas, emerging threats, opportunities, and alternatives, but is unable to sustain attention on any issue or approach that interests consumers. As a result producers neither develop sufficient expertise to understand current events, nor provide consumers with relevant information. By comparison, if analysts only devote themselves to those issues that are known to interest policy makers, then they will fail to detect and warn of emerging threats and opportunities. Therefore, strategic surprises can be framed as organizational failures that occur due to satisficing and not devoting enough time, attention, and resources to exploring the unknown, replicating the pathologies found amongst individual analysts in confronting their mindsets.

From Structured Analytic Techniques to Model-Centric Tradecraft

SATs were developed in order to address analysts' cognitive biases, and have transformed the negative connotations of mindsets into a neutral articulation and evaluation of mental models. By rendering analysis transparent, and explicitly providing consumers with the evidence used to justify particularly inferences, SATs not only improve the quality of analysis, but also its acceptability to consumers. While they have proven difficult to institutionalize, SATs have become the centerpiece of contemporary analytic tradecraft.

Although these manual methods have given renewed rigor in the production of intelligence, they can be improved by the inclusion of formal, computational modeling and simulation. Efforts to include modeling and simulation in the panoply of SATs have been difficult, however. Models, simulations, and games were included in early publications on AA and again on discussions of SATs.²⁹⁴ However, the prominence of models, simulations, and games has greatly diminished in the more recent literature, and Heuer and Pherson's recent book on SATs considered these approaches as a distinct form of analytic techniques all together.²⁹⁵

The use of formal models in analytic tradecraft has not received the same treatment as other SATs for two primary reasons. First, models, simulations, and games are often heavyweight methods, requiring an investment of resources that are qualitatively distinct from other SATs.

Gaming, Modeling, and Simulation are among the more sophisticated techniques taught in a more advanced analytic methods course and usually require substantial commitments of analyst time and corporate resources.²⁹⁶

Therefore, modeling and simulation possess a different resource profile than other SATs, requiring greater time, access to methodologists or other specialists, and a sustained commitment to strategic intelligence.

²⁹⁴ For examples, see the Directorate of Intelligence's 25-page *Alternative Analysis Primer* published in the mid-2000s, and its more recent *Structured Analytic Techniques: A Primer*, published in 2009. See Roger Z. George, "Fixing the Problem of Analytical Mind-Sets: Alternative Analysis," *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), p. 404; and Central Intelligence Agency, *Structured Analytic Techniques: A Primer* (Washington, DC: Central Intelligence Agency, 2009).

²⁹⁵ Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly, 2011), pp. 22-24.

²⁹⁶ Central Intelligence Agency, *Structured Analytic Techniques: A Primer* (Washington, DC: Central Intelligence Agency, 2009), p. 38.

A second difference between SATs and modeling and simulation is the character of their use. SATs allow for a direct assessment of analysts' mental models in order to produce qualitative insights into intelligence problems.²⁹⁷ Their emphasis is on the organization of assumptions, data, and uncertainties into formats that allow for individual analysts, their peers, managers, and consumers to identify, challenge, and communicate the aspects of the problem that have motivated analytic judgments.

Alternatively, formal models allow analysts to generate and explore externalized representations of the systems as an independent artifact. Thus, they blend substantive knowledge about intelligence issues with a means for exploring mental models. In these circumstances, formal models become engines for generating data upon which inferences can be based while simultaneously exposing and evaluating the mental models analysts used to create the model in the first place. How analysts choose to design the model, the variables they include (or exclude), the data or estimates used to set model parameters, their choices regarding experimental design, etc. simultaneously reveal the structure of mental models and show how systems would behave according to the structure and rules embedded within the formal model. By encouraging the use of formal modeling in analytic tradecraft, analysts can reap many of the benefits of more conventional SATs, aligning intelligence analysis with scientific practice, and improve the rigor of analytic production.

SATs were developed to aid analysts in recognizing when they were overly invested in particular mental models and satisficing in their search for alternatives. These

²⁹⁷ Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011), p. xvii.

techniques are manual, problem agnostic, and internally focused in the sense that they emphasize the explication and examination of mental models, but do not require substantive expertise on particular problems. Therefore, SATs mark a middle ground along the way toward the development of a model-centric analytic tradecraft. They admit that analysis is a type of modeling, and employ models as cognitive aids, but lack the use of models as a mechanism for providing substantive insights into particular policy problems.

Chapter 4: Intelligence and the Philosophy of Science

One of the ongoing debates within the intelligence community has been whether intelligence should be regarded as an art or a science—with most scholars and practitioners concluding that it has more in common with the former than the latter.²⁹⁸

This chapter examines several of the prominent issues in this debate, and argues that the character of this question is misleading. By reframing intelligence analysis as a particular type of scientific inquiry, the possibilities provided by a model-centric analytic tradecraft take on new importance.

The basis for rejecting science as a model for the practice of intelligence has rested on a highly stylized vision of scientific research, disciplines, and methods that mischaracterize its actual practice. The frame of reference used for comparing analytic tradecraft to science has been disconnected from both the history and philosophy of science. As a result, comparisons between these two intellectual pursuits have not only obscured important concepts that could aid the intelligence community, but have also allowed research practices to diverge, limiting the extent to which intelligence analysts have capitalized on emerging trends and practices in scientific communities.

²⁹⁸ Walter Lacquer, *A World of Secrets: The Uses and Limits of Intelligence* (New York, NY: Basic Books, 1985), pp. 293-308; and Peter Gill and Mark Phythian, *Intelligence in an Insecure World* (Malden, MA: Polity, 2008), pp. 1-8.

Although the alignment between intelligence and science can be improved, the intelligence community must choose carefully what parts of science it uses as a model for its own tradecraft. Given the diversity of sciences, the intelligence community has much to choose from. Therefore, intelligence scholars and practitioners have focused on the wrong question: rather than ask whether intelligence analysis is an art or science, more productive answers will come from asking what kind of science intelligence is or could be.

Intelligence and the Philosophy of Science

Several reasons have been offered to explain why intelligence does not conform to scientific practices. In a recent discussion of whether intelligence is an art or science, Marrin noted six reasons as to why intelligence and science differed:

...intelligence analysis may be even more uncertain than other forms of social science due to (1) externally imposed and frequently short deadlines which necessitate analysis prior to the acquisition of sufficient information, (2) an inability to control the variables under study, (3) an unknown data quality due to imperfections in the collection process, (4) the possibility of deception, (5) an emphasis on prediction, and (6) a focus on utility for the decisionmaker. Perhaps the most difficult methodological hurdle to overcome is the reality that intelligence analysts deal with information that is essentially a biased sample acquired opportunistically rather than according to some master research design, with frequently no way to determine how representative that sample is relative to the broader population.

In the end, the process of intelligence analysis has significant limitations that prevent many practitioners from wholeheartedly supporting the “intelligence analysis as science” perspective in the debate.²⁹⁹

²⁹⁹ Stephen Marrin, “Is Intelligence Analysis an Art or a Science?” *International Journal of Intelligence and Counterintelligence*, Vol. 25, No. 3 (2012), pp. 532-533.

These six reasons for discounting whether intelligence analysis can be practiced as a science vary considerably, but collectively capture many of the reasons why the intelligence community has often resisted efforts to impose scientific standards and methods on the training and practice of analysis.

The belief that intelligence analysis differs from science is not a universally held view, and several notable works have emphasized its practice as a scientific discipline or potential to become one. Sherman Kent emphasized its scholarly foundations and hoped that the community would eventually develop in a professional discipline complete with a guiding theory, rigorous training, and specialized literature.³⁰⁰ Richards Heuer, Jack Davis, Douglas MacEachin, and others examined psychological and cognitive processes that affect analysis, and advocated the development of tradecraft that increased the transparency, rigor, and considerations of analytic products.³⁰¹ Simultaneously, Ben-Israel examined the logic and practice of intelligence estimates from the perspective of the philosophy of science, and even engaged in an extended dialogue with Paul

³⁰⁰ Sherman Kent, *Strategic Intelligence for American World Policy* (Princeton, NJ: Princeton University Press, 1949); and Donald P. Steury, ed., *Sherman Kent and the Board of National Estimates: Collected Papers* (Washington, DC: Central Intelligence Agency, 1994).

³⁰¹ Richards J. Heuer, Jr., ed., *Quantitative Approaches to Political Intelligence: The CIA Experience* (Boulder, CO: Westview Press, 1978); Douglas J. MacEachin, "The Tradecraft of Analysis," in Roy Godson, Ernest R. May and Gary Schmitt, eds., *U.S. Intelligence at the Crossroads Agendas for Reform* (Washington, DC: Brassey's, 1995), pp. 63-75; Jack Davis, *A Compendium of Tradecraft Notes* (Washington, DC: Central Intelligence Agency, February 1997); Richards J. Heuer, Jr., *Psychology of Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 1999); Jack Davis, "Why Bad Things Happen to Good Analysts," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 157-170; Richards J. Heuer, Jr., "Computer-Aided Analysis of Competing Hypotheses," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 251-265; and Richards J. Heuer, Jr., and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly, 2011).

Feyerabend, a prominent philosopher, on the topic.³⁰² More recently, Bruce forcefully advocated intelligence analysts to adopt the practices of scientific discourse and investigation in order to improve analytic tradecraft and products.³⁰³ Specifically, Bruce argued that amongst the many alternative ways of developing knowledge, science was unique due to its self-corrective capabilities.

To summarize, what authority, habit of thought, rationalism, and empiricism all have in common is a demonstrated capacity for producing error as well as truth. But none of the four has the internal capacity to discover when it is wrong or to prescribe the needed correctives for getting it right. In a historically profound development that combines the third and fourth ways of knowing—rationalism and empiricism—the emergence of science produced a new epistemology that presents a powerful new feature to knowledge building: *self-corrective techniques*. Though all five avenues to knowledge can produce error, *only science has the built-in capacity to identify and correct its mistakes*. The implications for intelligence analysis are obvious and irresistible: These self-corrective techniques can markedly reduce the potential for error in analysis and greatly enhance the production of reliable knowledge.³⁰⁴

Collectively, these authors, Ben-Israel and Bruce most prominently, have addressed the arguments regarding the need to support policy makers (often on short time deadlines), and the problem of experimental control, deception, and prediction. These problems, as well as two additional challenges to the scientific practice of intelligence—secrecy and complexity—are discussed below.

³⁰² Isaac Ben-Israel, "Philosophy and Methodology of Intelligence: The Logic of Estimate Process," *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 660-718; and Isaac Ben-Israel, "Philosophy and Methodology of Military Intelligence: Correspondence with Paul Feyerabend," *Philosophia*, Vol. 28, No. 1-4, (June 2001), pp. 71-102.

³⁰³ James B. Bruce, "Making Analysis More Reliable: Why Epistemology Matters to Intelligence," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 171-190.

³⁰⁴ James B. Bruce, "Making Analysis More Reliable: Why Epistemology Matters to Intelligence," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 177.

Policy Support and the Problems of Relevance and Timeliness

The problem of policy support exists in two forms. The first is the production of information that is relevant to the needs of consumers, while the second is the need to provide this information to consumers in time for them to use it. As noted earlier, institutional and cultural differences between intelligence analysts and policy makers greatly affect how each defines what it means to support policy. The difficulties of this relationship, particularly within the context of the US intelligence community's professional standard of policy neutrality, was discussed at length by Sims, who concluded that producer/consumer relationship was a primary determinant of intelligence being an art, not science.

... perhaps the most difficult and unusual aspect of intelligence analysis is the analyst's responsibility to influence decision-makers without conveying any preference concerning the choices available. With this objective, intelligence analysis, when compared with other types, borders on the bizarre. This is because most analysts do have a sense of which choice will more likely gain advantages for their own side. Moreover, they develop a stake in their methodologies which, if successful in explaining past outcomes, shape expectations about future ones. This is no less true for intelligence experts. If, then, the policymaker chooses a course that is at odds with these expectations, how does an intelligence analyst not conclude the policymaker is wrong?...

That more restraint must be exercised in the modern U.S. system should not, however, prevent the intelligence analyst from caring about his own side's success or the outcomes of the policies they are supposed to inform. Successful intelligence provides advantages to decision-makers they would not otherwise have, so an analyst must know the decision-maker's frame of mind and strategy well enough to help the policymaker succeed. Intelligence forges a relationship of trust between partners seeking wins for their team. Thus, good intelligence is both objective and subjective and herein lies the essence of the analyst's conundrum: to be an expert and critical thinker, targeted for manipulation, legally denied

relevant knowledge, responsible for advising, but prohibited from judging. Maneuvering through this terrain is more than science; it is art.³⁰⁵

Sims' assessment was specific to the US intelligence community, however, and may therefore deny intelligence analysis the status of science due to idiosyncratic artifacts of the institutional relationships embodied in the structure of the US government. Therefore, it is also worthy to note Pillar's more generalized institutional perspective on intelligence and science. He noted that because intelligence analysts must support policy makers in their real-time decision making, they cannot wait for events to unfold before assessing contemporary issues and potential futures. As a result, while analysts must always confront the prospect of "intelligence failure," the scientific community does not face the same threat of "scholarly failure"—significantly affecting how each goes about their work.³⁰⁶ While science is predicated on transparency and intelligence rests on secrecy, within each of these communities relations are often reversed. Intelligence analysts must be transparent and cannot withhold their judgments from policy makers until they believe they have a firmer understanding of the situation and a stronger basis upon which to justify their assessments. Even though the finished products of intelligence are closely held, analysts actually work in a transparent fashion in anticipation of constant disruptions and interruptions from consumers who may demand analytic support before careful, measured final judgments are available. Alternatively,

³⁰⁵ Jennifer E. Sims, "Decision Advantage and the Nature of Intelligence Analysis," in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), pp. 390-392.

³⁰⁶ These comments were made by Paul Pillar during an interview with the author February 1, 2012. Also see Aaron B. Frank, *Interview with Paul Pillar from Georgetown University*, <http://www.aaronbfrank.com/2012/06/interview-with-paul-pillar-from-georgetown-university/> (accessed on June 19, 2012).

scientists can choose to not publish their data or findings until they believe it is ready to withstand the scrutiny of their peers. Thus, while scientists may produce publically available knowledge, their working processes may be opaque, privately held, and only released when completed.

While policy concerns drive the production of intelligence, their presence does not negate the opportunity for analysis to be performed scientifically. Ben-Israel provided a forceful critique of the argument that supporting policy makers' demands deviates from the practice of science by comparing intelligence failures with technology and the products of applied science.

A failed experiment—a refutation—can be a great success in science but catastrophic in intelligence. Such a failure is, however, disastrous in every part of the *technological* domain. A wrong prediction resulting in the building of an allegedly unsinkable ship, can lead to its sinking on its maiden voyage, drowning most of its passengers, as happened with the *Titanic*. There is no difference, from this point of view, between intelligence and ship-building or any other technology. The argument concerning the cost of failure is thus refuted by the technology example. The scientific method is as suited to intelligence research as it is to technology, despite fears of failures. The scientific method cannot prevent failures; but it can use them systematically for learning and improving its theories. So, if we cannot avoid new failures, the scientific method can, at least, prevent us repeating the *same kind* of failures—the only method that can do this.³⁰⁷

Characterizing Intelligence as an applied science, such as engineering, architecture, medicine, and management, emphasizes the development of useable knowledge that enables decision makers to shape the environment and events as desired. As Herbert Simon noted, these artificial sciences are all cases of a broader science of design.

³⁰⁷ Isaac Ben-Israel, "Philosophy and Methodology of Intelligence: The Logic of Estimate Process," *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 696.

The thesis is that certain phenomena are “artificial” in a very specific sense: they are as they are only because of a system’s being molded, by goals or purposes, to the environment in which it lives. If natural phenomena have an air of “necessity” about them in their subservience to natural law, artificial phenomena have an air of “contingency” in their malleability by environment.

The contingency of artificial phenomena has always created doubts as to whether they fall properly within the compass of science. Sometimes these doubts refer to the goal-directed character of artificial systems and the consequent difficulty of disentangling prescription from description. This seems to me not to be the real difficulty. The genuine problem is to show how empirical propositions can be made at all about systems that, given different circumstances, might be quite other than they are....

Finally, I thought I began to see in the problem of artificiality an explanation of the difficulty that has been experienced in filling engineering and other professions with empirical and theoretical substance distinct from the substance of their supporting sciences. Engineering, medicine, business, architecture, and painting are concerned not with the necessary but with the contingent—not with how things are but with how they might be—in short, with design. The possibility of creating a science or sciences of design is exactly as great as the possibility of creating any science of the artificial.³⁰⁸

That intelligence is subservient to policy makers does render it distinct from pure science, but not applied science. The demands of consumers may complicate its practice, but not in ways that are unfamiliar to many scientists who must develop new technologies on command, whether in the form of mechanical or electronic systems, management, organizational design, or legal code intended to produce desired results.³⁰⁹ However, the need to understand the character of contingency in intelligence and policy problems does challenge many of the epistemological underpinnings of science and its practice.

The challenge of timely intelligence production limits the ability to allow events to unfold and wait for more information. Moreover, many times pressures are

³⁰⁸ Herbert A. Simon, *The Sciences of the Artificial* (Cambridge, MA: MIT Press, 1996), pp. xi-xii.

³⁰⁹ William H. McNeill, *The Pursuit of Power* (Chicago, IL: University of Chicago Press, 1982).

institutionally driven, particularly as they relate to the management of analytic resources and attention regarding the production of current vs. strategic intelligence. Therefore, intelligence cannot fully behave according to the precepts of inductive science, where the observation adjudicates the validity of a given hypothesis. This chapter will address the more important challenge that time poses to the treatment of intelligence as science, especially as it relates to inductive inference.

Experimental Control and Data Collection Limitations

The lack of systematic control over variables or dependence on erratic, often uncontrolled sampling was also used to differentiate between intelligence as art or science. These critiques, however, are not cause for dismissing the notion of intelligence as science.

John Lewis Gaddis noted that many sciences are historical, and rely on the careful gathering of data, observation, and simulation of artificial worlds—whether in the form of informal mental models or formal models. Thus, modeling and simulation served as the primary laboratory in which historical processes and events were understood.

The key to consensus, in science, is reproducibility: observations made under equivalent conditions, no matter who makes them, are expected to produce closely corresponding results.... Verification, within these disciplines, takes place by repeating actual processes. Time and space are compressed and manipulated: history itself is in effect rerun. In that sense, obviously, the historical method can never approximate the scientific method.

But not all sciences work this way. In fields like astronomy, geology, paleontology, or evolutionary biology, phenomena rarely fit within laboratories, and the time required to see results can exceed the life spans of those who seek them. These disciplines instead depend upon thought experiments: practitioners rerun in their minds—or perhaps now in their computer simulations—what their test tubes, centrifuges, and

electron microscopes can't manage. They then look for evidence suggesting which of these mental exercises comes closest to explaining their physical observations. Reproducibility means building a consensus that such correspondences seem plausible. The only way these scientists can rerun history is to imagine it, but they must do so within the limits of logic. They can't attribute the inexplicable to pixies, wizards, or extraterrestrial visitors and still expect to persuade their peers that their findings are valid....

It's here that the methods of historians and scientists—at least those scientists for whom reproducibility cannot take place in the laboratory—roughly coincide. For historians too start with surviving structures, whether they be archives, artifacts, or even memories. They then deduce the processes that produced them. Like geologists and paleontologists, they must allow for the fact that most sources from the past don't survive, and that most daily events don't even generate a survivable record in the first place. Like biologists and astrophysicists, they must deal with ambiguous or even contradictory evidence. And like all scientists who work outside of laboratories, historians must use logic and imagination to overcome the resulting difficulties, their own equivalent of thought experiments, if you will.³¹⁰

Gaddis's point is evident when considering the fossil record and the challenges faced in evolutionary biology and paleontology, where determining the record's completeness and biases is a non-trivial process and consistent with the challenges faced by intelligence collectors and analysts.

The chances of an organism becoming fossilized are vanishingly small. The familiar hierarchical food chain and the biological cycles that sustain life are based on the breakdown and recycling of nutrients that dead organisms contain. To become a fossil, at least some part of the animal or plant must avoid decay and physical destruction, and with few exceptions, such as trapping or insects in resin exuded from trees (e.g., amber), it must be buried in sediment. Because most sedimentary deposition occurs in the sea, marine organisms have a much higher chance of becoming entombed and fossilized than the animals and plants that live on the land. Likewise organisms that live in lakes and rivers, or close to them, are better represented than animals and plants that live away from water. Burial, however, is only the beginning of the story—the sedimentary rocks that

³¹⁰ John Lewis Gaddis, *The Landscape of History: How Historians Map the Past* (New York, NY: Oxford University Press, 2002), pp. 39-41.

contain a potential fossil must survive geological processes such as the addition of overlying layers of sediment (with its attendant rise in pressure and temperature), uplift and erosion, and the destruction of ocean floor as tectonic plates descend into trenches at the margin of continents. Because of this last process there is no oceanic crust more than about 180 million years old flooring today's ocean. Older marine rocks that have survived were deposited in basins on continental crust or thrust onto the continents during episodes of mountain building. Finally the fossil must be found and described.³¹¹

Thus, problems of collection bias and a lack of control or confidence in the veracity of information sources resemble the same challenges intelligence analysts face when evaluating intelligence collection.

Again, Ben-Israel provides important insight. He noted that science may be driven by data, but whether that data is gathered under controlled, experimental conditions or through careful observation of the natural and social world is philosophically immaterial. The laboratory may facilitate scientific discovery and the testing of theory, but it is not a requirement.

The problem of uniqueness and the impossibility of *exactly* replicating the same experiment is as acute in medical and biological science as in social science. An experiment in which a laboratory animal is given an experimental drug and then examined *post mortem* can never be repeated with the same animal....

What is the meaning of "experiment" at the methodological level? We are not interested here in the *practical* aspects of experiments..., but in their *logical* function. From this point of view, an experiment is a logical confrontation between two statements, one of which is a *prediction* (deduced from a conjunction of theory—a set of "laws of nature"—and initial data) and the other an *observation report*. An experiment need not be active. Even a passively obtained observation report can be used to refute theoretical hypotheses. The best example here is, perhaps, astronomy. No one has ever been on a star outside the solar system and it is unlikely that any earthling ever will, as the nearest star is more than four

³¹¹ Nicholas H. Barton, Derek E. G. Briggs, Jonathan A. Eisen, David B. Goldstein, and Nipam H. Paten, *Evolution* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2007), pp. 253-254.

light years away. Despite this, astronomy is a “kosher” science, which for centuries was considered queen of all sciences. Yet all astronomy “experiments” are passive: we collect observation data which arrive at earth, but cannot conduct active experiments (in the atmosphere of Alpha-Centauri for example).

From the scientific methodological point of view, we do not need “experimental information,” only observational information. The so-called experiment is nothing more than a logical testing of predictions against observed information. Similar observed (“factual”) information exists in intelligence as well.

Indeed, when we are able to initiate an experiment, we can shorten the lengthy process of creating and refuting theories by planning a special experiment designed to produce falsifying information. But the same is possible in intelligence. Let us assume, for example, that we hypothesize that the enemy will not carry out action A unless condition B is fulfilled. Now, whenever the enemy takes action A (during training, for example), we can direct our gathering agencies to watch for B. This is a “kosher” test for our hypothesis. It may even be a crucial test should we have an alternative hypothesis which predicts the existence of A without B.³¹²

The critique regarding the distinct problem of working with intelligence sources as a problem distinct from science is equally problematic. While TECHINT, HUMINT, and OSINT have distinct biases and artifacts, these kinds of errors or limitations are not foreign to scientists working in the laboratory or the field. Indeed, many of the most sophisticated physical experiments are only possible through the use technological aids, each capable of generating their own artifacts, for example controlling and observing particles at desired spatial and temporal scales. The history of various physical constants, such as the W boson or the Hubble Constant, show how it may take decades for measurements and estimates to converge to consistent values, often far from their original

³¹² Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 685-686.

estimate, due to changes in experimental techniques and technology (see Figure 4-1 below).³¹³

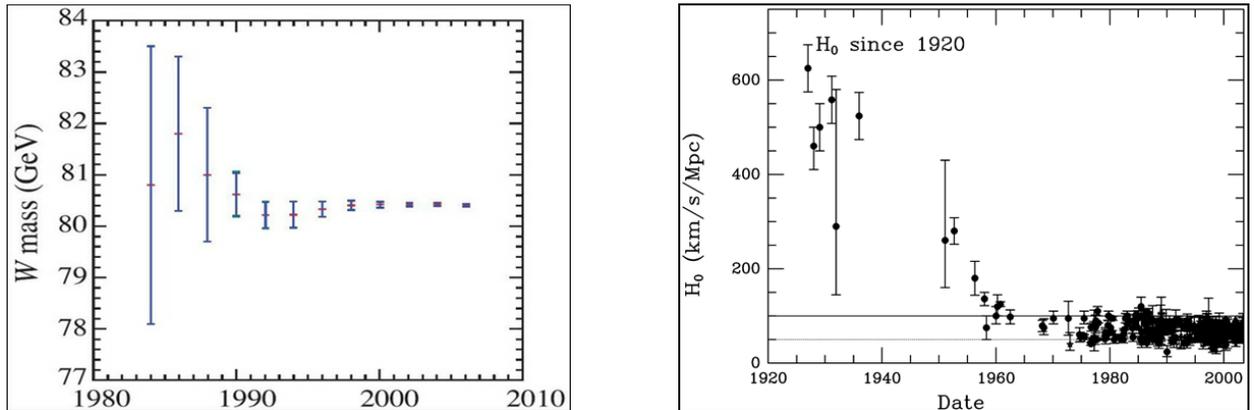


Figure 4-1: Measurement and Estimates of Physical Constants over Time. The figure on the left shows the estimate of the W boson over time based on repeated calculations. The figure on the right shows the estimate of Hubble Constant since first calculated in 1920. In both cases, measurements have largely converged over time, but not smoothly. Importantly, each constant showed periods where the estimated precision of calculations was reduced and the mean value was far removed from later calculations than prior, more uncertain calculations.

The fact that intelligence analysts depend on collection systems that have known and unknown biases and limitations is important with respect to how it effects their analysis, but is not an issue that scientists are unfamiliar with, and therefore not a credible distinction between science and intelligence.

³¹³ Troels Petersen, *Measuring the W Mass with ATLAS*, February 11, 2008, http://atlas-service-enews.web.cern.ch/atlas-service-enews/2007-8/features_07-8/features_wmass.php (accessed on June 20, 2012); and Ed Hawkins, *The Importance of Reliable Uncertainty Estimates*, January 3, 2012, <http://www.climate-lab-book.ac.uk/2012/reliable-uncertainty/> (accessed on June 20, 2012).

Deception

The challenge of deception is frequently considered a significant concern in intelligence and is often regarded as undermining scientific approaches to analysis by altering the relationships between evidence and inference. Many of the problems posed by deception are logical extensions of the concerns over data collection where adversaries seek to manipulate what information is collected by rivals and how it is interpreted. For example, Robert Jervis noted social scientists rarely have to worry about the manipulation of their information by the subjects of their research.

In many cases, we are trying to predict the actions of people who are, or may be, trying to mislead us... suffice it to say here that the possibility of deception renders invalid and complete analogy between intelligence and social science, let alone physical science. Social scientists rarely have to worry about more than the danger that those they are studying are trying to conceal important facts from them; nations often actively try to mislead one another. The use of “turned” agents is only the most dramatic illustration of this problem. Indeed, if one country discovers what indices or aspects of behavior the other is using to draw inferences, it may be able to manipulate them in order to project a desired, although often misleading, image. Furthermore, the knowledge that the other may be attempting deception must lead intelligence analysts to discount various sources of information that in fact may be reliable.³¹⁴

Likewise, Sims argued that the problem of deception was one of the primary reasons for deviating from more traditional scientific practices, in spite of the fact that the remainder of their research, communication, and analytic skills parallel those from other disciplines.

In these skills, and in those of communicating effectively through the written and spoken word, sophisticated analysts are not that different from each other, whether employed in intelligence, academia, gambling, business, politics, or medicine.

³¹⁴ Robert Jervis, “Improving the Intelligence Process: Informal Norms and Incentives,” in Alfred C. Mauer, Marion D. Tunstall, and James M. Keagle, eds., *Intelligence: Policy and Process* (Boulder, CO: Westview Press, 1985), p. 114.

Yet to suggest that the job of an intelligence analyst stops here, with the tasks of critical thinking and clear writing, is to overlook the obvious and most difficult aspects of the work: secrecy, urgency, deception, and influence.... For these and other reasons, intelligence analysts must be trained to recognize and cope with deception—that is, deliberate efforts by adversaries to frustrate, taint, or disrupt their work. Here, intelligence analysis becomes tradecraft. After all, a scientist rarely analyzes specimens that lie not only *in* the Petri dish but *from* it as well.³¹⁵

In contradiction to Jervis and Sims, others in the intelligence community have argued that grounding analytic tradecraft in scientific practice is essential for detecting and coping with Denial and Deception (D&D). Bruce and Michael Bennett noted that the constant testing and evaluation of mental models using SATs were essential to the analytic process, adopting practices that deliberately encourage intelligence analysts to emulate scientific methods of model development and hypothesis testing.

Our desired goal is to deliver better, more accurate judgments that will enable the negating or at least mitigation of the effects of denial and deception. The first broad strategy for achieving the goal of reducing the mind's vulnerability to D&D is to improve the analytical process. The *know yourself*, *know your adversary*, and *know your situation* principles highlight the importance of two interdependent approaches: mitigating cognitive biases, and adopting systematic or “structured” methodologies....

The *know yourself* principle emphasizes recognizing the assumptions, preconceptions, and expectations that influence analyst beliefs, while the *know your situation principle* focuses on continually evaluating the environment for the cues that deception may be a factor in the situation under consideration. The use of structured analytical methodologies or “challenge” analysis also provides another way of restructuring problems so that assumptions, preconceptions, and mental models—that is, factors shaping mindsets—are not hidden by making them more explicit so that they can be examined and tested. In particular, such structured methodologies include Analysis of Competing Hypotheses (ACH), argument mapping, and signpost analysis; “challenge analysis” techniques include Devil's Advocacy, What-If Analysis, and High-

³¹⁵ Jennifer E. Sims, “Decision Advantage and the Nature of Intelligence Analysis,” in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), p. 390.

Impact/Low-Probability Analysis. Using such methodologies can reduce the likelihood that important biases or situational cues are not recognized or ignored.

A prepared mind will make a conscientious effort to see the problem or situation from the adversary's point of view. It will continually test and retest its judgments, update and evaluate all the evidence at hand, and remain alert to cues and anomalies in the environment that something has changed or is missing. It will not ignore its intuition when something does not quite feel right about a complex analytical situation. And it will diligently update and evaluate the credibility of information sources, stay alert to any channels that may have been compromised, and revisit the issue of source vetting and validation.³¹⁶

Within the social sciences, the challenges posed by deception are familiar to researchers. Psychologists, economists, and sociologists are all well aware that experimental and observational subjects may engage in deception. Social scientists often discuss the Hawthorne effect, which notes that people's behavior changes when they are aware that they are being observed, endowing the act of data gathering with causal powers. As Godfrey-Smith noted, observers may exert causal power over the behavior of the observed.

Suppose you want to know how many teenagers smoke. The obvious way to answer this question is to collect a random sample of teenagers, find the rate of smoking in the sample, and extrapolate to the larger teenage population, in a way guided by statistical measures of likely error. Why should we do this rather than something else? Why not find the proportion in the sample and extrapolate half that proportion, or one minus that proportion? (This second option would be a kind of counter-induction.) *Not* because of an equilibrium between intuitions—or at least, not at this stage in the analysis. Instead, we have a statistical model of why the procedure is in principle a reliable one. The model tells us how samples of different sizes will be distributed, in relation to the actual properties of the population being sampled. It tells us when, and the

³¹⁶ James B. Bruce and Michael Bennett, "Foreign Denial and Deception: Analytical Imperatives," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 133-134.

extent to which, the properties of a sample are reliable indicators of the properties of the underlying population.

I said that the question about teen smokers could be answered using inference from a sample. But there are various ways this method might fail. Maybe you cannot collect a random sample, as the smokers tend to avoid you. Perhaps teenagers will not tell you the truth. There is also a third, more unlikely possibility. Perhaps being asked the question tends to *make* some teenagers immediately take up smoking. So they truthfully answer that they smoke, but only because they were asked. The process of data-gathering is interfering with the objects you are observing, in a way that makes them an unreliable guide to the unobserved cases. This is not a case of “selection bias”; we can assume that the original collection of teenagers asked about their smoking really was a random sample. Some statisticians call this a “Hawthorne effect,” after a famous case in the 1930s. It also has a kinship with the notion of a “confounding variable,” although that term is usually applied in the context of causal inference rather than estimation. The term “observation selection effect” is used in some literature to cover various phenomena, including sample bias, the confounding-like phenomenon discussed here, and maybe others. In any case, that special relationship between surveying and smoking would make the method fail. The problem has nothing to do with “non-projectible” predicates; it has to do with the process of collecting our sample and some unwelcome causal relations.³¹⁷

Importantly, the Hawthorne effect is so powerful that it undergirds the basis of surveillance as a form of political power due to an individual’s fear of being observed and identified acting in violation of the law and norms, thus deterring criminal and undesired behavior.³¹⁸

Ben-Israel also addressed the problem of deception and the ways in which the possibility of being observed can alter the behavior of targets by comparing intelligence collection and social science with physics and Heisenberg’s uncertainty principle.

³¹⁷ Peter Godfrey-Smith, “Induction, Samples, and Kinds,” in Joseph Keim Campbell, Michael O’Rourke and Matthew H. Slater, eds., *Carving Nature at Its Joints: Natural Kinds in Metaphysics and Science* (Cambridge, MA: MIT Press, 2011), pp. 37-38.

³¹⁸ James Der Derian, “Anti-Diplomacy, Intelligence Theory and Surveillance Practice,” *Intelligence and National Security*, Vol. 8, No. 3 (July 1993), pp. 40-44.

A[n]...argument against the applicability of scientific method to intelligence, concerning the concept of “experiment,” is based on the (true) observation that *social experiments affect the predicted results*. Society, it is argued, unlike nature, has history, memory and ability to learn. Hence, every action taken in a society (e.g., an “experiment”) may have a totally different result from what would have happened had the experiment not taken place. Let us assume (counter-factually) that Israel’s official estimate before the Yom-Kippur War was that war was imminent and, as a result, underwent general mobilization. The possibility cannot be ruled out that this mobilization alone could have caused the Arabs to postpone their plans—and thus result in an apparent failure of the estimate!

The argument about the experiment affecting its outcome is not confined to intelligence. It is valid in science as well, where it is well known that a scientific measurement affects the measured system (Heisenberg’s uncertainty relations). And while the change caused by measurement in physics is small and confined to the sub-atomic realm and that in intelligence can have macroscopic results, the difference is strictly quantitative. It has no implications on the methodological level. Essentially, a complex relationship between the “measurement” and the “measured” *can* be treated within the scientific method.³¹⁹

Once again, the prospect of D&D complicates the work of intelligence analysis, but its occurrence does not invalidate scientific approaches to intelligence.

Prediction

Prediction poses a challenge to intelligence analysts, but its place in science is just as problematic. Marrin noted that one of the differences between science and intelligence is the intelligence community’s emphasis on prediction, but prediction itself is not a settled issue within the community. As noted earlier, consumers often resist predictions from the intelligence community, particularly those that differ from their own expectations and are detached from any causal claims by which the emergence of threats and opportunities can be explained.

³¹⁹ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 686.

Arguments regarding prediction in intelligence and its comparison with science are more often indications of defects in producer/consumer relations and a misunderstanding of scientific practice, history, and philosophy than a serious challenge to the epistemological foundations or purpose of tradecraft.

In addition to the six challenges to viewing intelligence as a science identified by Marrin, two additional issues have often received attention: the consequences of institutionalized secrecy and the problem of analyzing open, complex systems.

Secrecy and Complexity

Science is often regarded as thriving in open societies, where ideas and issues can be freely debated and challenged. As a result, the secretive character of intelligence collection and analysis has been seen as antithetical to scientific practice. For example, Daniel Moynihan argued that the employment of secret sources and security measures used in intelligence distorted the normal practices of science and left the intelligence community and the policy makers they supported worse off as a result.³²⁰

Ben-Israel considered this issue in his own exploration of science and its applications to intelligence, particularly with respect to its deviation from Popper's ideal of an open society where authority is regularly challenged.

Another difficulty in applying the scientific method is rooted in the making of intelligence estimates within a hierarchical organization. Civil intelligence agencies, like the CIA, are also hierarchically organized with ranks, military discipline, and so on. In addition, the secrecy which is vital to intelligence work tends to limit severely the internal circulation of intelligence material, and may result in information being concealed from intelligence researchers themselves.

³²⁰ Daniel Patrick Moynihan, *Secrecy* (New Haven, CT: Yale University Press, 1998).

Consequently, the intelligence community is a *closed society*—a society in which information is not easily obtainable, in which freedom of speech is limited, and in which effective criticism thus scarcely exists. The intelligence community is, in fact, a paradigm of a closed society. As indicated earlier, an open society is a necessary precondition for the fruitful application of the critical method. As this is not the case in intelligence, it seems that application of the scientific method is impractical.³²¹

Ultimately, Ben-Israel conceded that secrecy posed a serious challenge, but not one beyond remedy.

The fact that the intelligence community is a closed society, without either freedom of information or freedom of criticism, is potentially a major obstacle to the development of intelligence theory. How can this problem be resolved? I doubt that it can be entirely resolved. Classified material and security of sources are, after all, real problems, but their vicious influence can be reduced....

In other words, while intelligence estimation cannot be opened to the general public, the micro-climate of an open society can be created within the intelligence community itself.³²²

The challenge posed by complexity is among the most serious faced by the intelligence community because it requires a holistic analytic approach, which has proven difficult for science to define and implement. For example, Ben-Israel was deeply skeptical of holism, arguing that it was not a scientific possibility.

The starting point of anti-scientific historicism is the view that *an enemy equals more than the sum of its constituent individuals*. An enemy (or, in general, a society) resembles a *biological organism* more than it does a physical system (which is fully determined by its constituent elementary particles). Like a biological organism, an enemy has history and memory, and can learn and adjust his behavior to a varying reality. All the factors influencing its behavior are integrally connected and cannot be isolated. Hence, the argument goes, we must investigate the enemy as a *whole*. It is never enough to know only its language, structure, economy etc. Its

³²¹ Isaac Ben-Israel, "Philosophy and Methodology of Intelligence: The Logic of Estimate Process," *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 690-691.

³²² Isaac Ben-Israel, "Philosophy and Methodology of Intelligence: The Logic of Estimate Process," *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 695.

undivided totality must be grasped in order to predict its behavior. Thus, the appropriate intelligence method must be *holistic*...

Holism can lead to the demand of conducting holistic “experiments” acting on the whole researched society. Experiments like these are either impossible, or inappropriate for falsifying hypotheses (if they act on the “whole” society, how can we know what specific hypothesis is refuted?).³²³

However, Ben-Israel later suggested that holism was possible when systemic understandings were generated from methodologically individualist approaches, such as focusing on the capabilities, intentions, actions, and interactions of a system’s units, rather than appeals to the system’s metaphysical essence.

Holism and the demand for understanding “essences” are typical of a “pre-scientific” or primitive stage of science. The so-called scientific revolution of the seventeenth century owes much of its enormous progress to its abandonment of understanding essences in favour of *informative explanations*. Trying to answer “questions of essence” (What is the enemy? What is the essence of Syrian rule?) will not take us anywhere. Instead, we should construct nominalist models, studying the interests, expectations, mutual relationships and *modus operandi* of relevant individuals. This methodological individualism is far more fruitful than the effort to “understand” essences.³²⁴

Ben-Israel’s perspective on holism mirrored debates over emergent properties and reductionism in the study of complex systems, particularly whether or not methodologically individualist approaches to social systems could adequately explain their behavior. In an argument that paralleled the Ben-Israel’s logic as to what constituted and what did not constitute scientific study, Joshua Epstein argued that ABM, one of the primary research tools for studying complex adaptive systems, was consistent with established scientific practice:

³²³ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 698.

³²⁴ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 701.

Typical of classical emergentism would be the claim: *No description of the individual bee can ever explain the emergent phenomenon of the hive.* How would one know that? Is this a falsifiable empirical claim, or something that seems true because of a lax definition of terms? Perhaps the latter. The mischievous piece of the formulation is the phrase “description of the individual bee.” What is that? Does “the bee’s” description not include its rules for interacting with other bees? Certainly, it makes little sense to speak of a Joshua Epstein devoid of all relationships with family, friends, colleagues, and so forth. “Man is a social animal,” quoth Aristotle. My “rules of social interaction” are, in part, what make me me. And, likewise, the bee’s interaction rules are what make it a bee—and not a lump. When (as a designer of agent objects) you get these rules right—when you get “the individual bee” right—you get the hive, too. Indeed, from an operationist viewpoint, “the bee” might be defined as that x which, when put together with other x ’s, *makes* the hive (the “emergent entity”). Unless the theoretical (model) bees generate the hive when you put a bunch of them together, you haven’t described “the bee” adequately. Thus, contrary to the opening emergentist claim, *it is precisely the adequate description of “the individual bee” that explains the hive....*

Classical emergentism holds that the parts (the microspecification) cannot explain the whole (the macrostructure), while to the agent-based modeler, it is precisely the generative sufficiency of the parts (the microspecification) that constitutes the whole’s explanation!³²⁵

The challenge of complexity in intelligence may be an enduring one, but again, is not qualitatively distinct from science. In fact, many of the major developments in the sciences have been the increasing emphasis on interdisciplinary research and methods for examining non-linear, complex systems that the intelligence community has always focused their collection and analytic efforts on. Thus, rather than limit the extent to which intelligence resembles science, the challenge of complexity is actually bringing the two communities closer with respect to the character of the questions they ask and the means available for their study.

³²⁵ Joshua M. Epstein, *Generative Social Science: Studies in Agent-Based Computational Modeling* (Princeton, NJ: Princeton University Press, 2006), pp. 36-37.

Upon reflection it is clear that each of these issues presents significant challenges to the practice of intelligence as science, but none are unique to intelligence or foreign to the experience of scientists. In some cases, a better appreciation for the history of science would reveal that distinctions between art and science are largely of quantitative degree rather than qualitative difference. In other cases, replicating the practices of scientific disciplines such as physics and chemistry may prove problematic for intelligence analysts, while the social sciences—ecology, biology, and applied sciences such as engineering and technology—may provide profitable insights.

What is Science?

Before examining how intelligence can be scientific, it is necessary to consider what it means to *be scientific*. Philosophers have tried to answer this question in many ways, most famously by trying to differentiate science from other methods of pursuing knowledge, such as theology. Thus, definitions of science are intimately related to the criteria for demarcating science from non-science, pseudo-science, or scientism.³²⁶

As has been the case in intelligence studies, where practitioners have often questioned the need for theory, scientists have often been dismissive of the philosophy of science. For example, Richard Feynman famously stated that, “Philosophy of science is about as useful to scientists as ornithology is to birds.”³²⁷ Likewise, John Ziman, a physicist turned philosopher, noted:

³²⁶ E. D. Klemke, “What is Philosophy of Science?” in E. D. Klemke, Robert Hollinger, and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), pp. 19-37.

³²⁷ Online Philosophy Club, *Philosophy of Science*, <http://onlinephilosophyclub.com/philosophy-of-science.php> (accessed on January 26, 2012).

One can be zealous for science, and a splendidly successful research worker, without pretending to a clear and certain notion of what science really is. In practice it does not seem to matter.³²⁸

Even prominent philosophers have questioned their field's contributions to the practice of science. For example, during an extended exchange of letters with Ben-Israel, Paul Feyerabend noted:

... while the philosophical ideas that affected the sciences in the past were closely connected with scientific practice and shared its fruitful imprecision, the ideas that come from modern philosophy of science (up to and including Popper and to some extent even Kuhn) are part of a school philosophy that gives some general and very misleading outlines but never descends to details....

All they did was to give historically incorrect accounts of the origin of relativity (here see the wonderful article by Holton on Einstein and the Michelson experiment), of quantum theory and so to confuse people instead of helping them.³²⁹

However, just as in intelligence, where theory serves as a backstop to practice, aiding it moments of uncertainty and crisis, philosophy assists scientists by evaluating the experimental and inferential practices embedded in their individual and collective work. In this role, philosophy does more to determine the extent to which scientific claims about knowledge and truth are accepted as universal or contingent, rather than advance any particular field through new discoveries and inventions. Therefore, philosophy is less of a guide that demarcates which paths researchers should follow, than markers that establish the boundaries within which science operates as the free exchange, pursuit, and

³²⁸ John Ziman, "What is Science?" in E. D. Klemke, Robert Hollinger, and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), p. 49.

³²⁹ Paul Feyerabend in Isaac Ben-Israel, "Philosophy and Methodology of Military Intelligence: Correspondence with Paul Feyerabend," *Philosophia*, Vol. 28, Nos. 1-4, pp. 79-80.

testing of ideas, and outside of which the development of knowledge becomes subjected to other forces.

The principal task of philosophy of science is to analyse the methods of enquiry used in the various sciences. You may wonder why this task should fall to philosophers, rather than to the scientists themselves. This is a good question. Part of the answer is that looking at science from a philosophical perspective allows us to probe deeper—to uncover assumptions that are implicit in scientific practice, but which scientists do not explicitly discuss. To illustrate, consider scientific experimentation. Suppose a scientist does an experiment and gets a particular result. He repeats the experiment a few times and keeps getting the same result. After that he will probably stop, confident that were he to keep repeating the experiment, under exactly the same conditions, he would continue to get the same result. This assumption may seem obvious, but as philosophers we want to question it. Why assume that future repetitions of the experiment will yield the same result? How do we know this is true? The scientist is unlikely to spend too much time puzzling over these somewhat curious questions: he probably has better things to do.³³⁰

The philosophy of science, much like the theory of intelligence, places what it means to be scientific in context. While practicing scientists often resort to doing what they know, philosophers and philosophically minded scientists regularly seek a wider perspective. These perspectives extend beyond viewing science as a matter of data collection, theorization, and hypothesis testing by considering the social context in which scientific pursuits occur and implement the scientific method.³³¹ As a result, alternative perspectives on the philosophy of science provide more than guidance on analytic matters such as how evidence and inference should be handled, but also speak to whether the design of the analytic production processes, the organization of intelligence community,

³³⁰ Samir Okasha, *Philosophy of Science: A Very Short Introduction* (New York, NY: Oxford University Press, 2002), p. 12.

³³¹ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), p. 12.

and the institutional arrangements between producers and consumers encourage or discourage scientific behavior.

Perspectives on Demarcation

One of the most fundamental questions in the philosophy of science is perhaps the most basic: What is the definition of science? This question has earned its own special name called the problem of *demarcation*, which provides answers regarding how to differentiate scientific from non-scientific practices and ideas. The debates over demarcation are relevant to intelligence analysis for four reasons, each touching on different aspects of intelligence analysis and tradecraft, analytic training, the organization of the intelligence community, and producer/consumer relations.

First, the existence of the debate over demarcation, and its lack of resolution, reflects the complexity of science as a method for developing and evaluating knowledge. The intelligence community's difficulties with these same definitional issues have been well documented, and have generally been seen as a sign of the field's immaturity when compared with other social sciences or professional disciplines. While the lack of clear definition of intelligence has been seen as a detriment, it may be that intelligence has more in common with science in general, and the philosophy of it, than with specific scientific disciplines whose purposes, processes, and goals are stationary or change slowly. Given the need for the intelligence community to operate on the frontiers of what is known, knowable, and relevant to the ever-changing needs of policy makers, any definition of what intelligence is, is not, and how it is practiced will be fleeting and prone to revision, just as has been the case in science.

Second, definitions of science have been multi-scale, emphasizing the role of ideas, the individual scientist, scientific institutions, and entire communities. The shifting focus of demarcation criteria that change emphasis from the individual scientist to more complicated mixtures of groups, social practices, history, and beliefs has mirrored the study of intelligence failures, where group processes, organizational design, and producer/consumer relationships have been given greater emphasis to be placed on par with individual beliefs, logic, data, and the act of “dot connection.”³³² Just as intelligence failures are increasingly viewed as complex events that cannot be reduced to a mistake by a single analyst or collector, definitions of science have extended beyond the merits of particular ideas or the activities of singular scientists in order to consider tradeoffs between cooperation and competition between individuals and groups in their collective pursuit of new knowledge.³³³

Third, definitions of science have clashed based on whether they have been interpreted as factual, empirical descriptions of how scientists behave, or normative arguments regarding how they should operate. This parallels intelligence studies where descriptions of tradecraft and methodology often result in normative prescriptions that are not followed, and can even prove problematic or counterproductive in practice.³³⁴

³³² Mark M. Lowenthal, “Towards a Reasonable Standard for Analysis: How Right, How Often on Which Issues?” *Intelligence and National Security*, Vol. 23, No. 3 (2008), pp. 303-315.

³³³ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003).

³³⁴ Roger Z. George, “Fixing the Problem of Analytic Mind-Sets: Alternative Analysis,” *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), pp. 399-401; and Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York, NY: Columbia University Press, 2007), pp. 42-43.

Fourth, the problem of demarcation reveals science to be a dynamic process that encapsulates the conduct of research, discovery, training, and changes in beliefs overtime. Indeed, many of demarcation criteria advanced by philosophers are less concerned with the logic of specific arguments or theories, than with how scientists adapt their beliefs and practices when they encounter new facts and theories. The ways in which scientists explore the unknown and evaluate and change their beliefs provide a template for thinking about the dynamics of intelligence analysis and tradecraft.

The following examination of the demarcation criteria shows that many of the intelligence community's ideas about science that have largely bolstered arguments as to why intelligence cannot be practiced scientifically, are in fact weak and often misleading when put into a larger philosophical context. The demarcation criteria of Karl Popper, Ziman, Thomas Kuhn, Imre Lakatos, and Paul Thagard are discussed below.³³⁵

Karl Popper

Popper's demarcation criterion set the foundation for debating the problem for decades. Popper argued that science was not about the discovery of truth, but rather a system for determining the merits of particular claims or conjectures about the world.

The problem which troubled me at the time was neither, "When is a theory true?" nor, "When is a theory acceptable?" My problem was different. I *wished to distinguish between science and pseudo-science*; knowing very well that science often errs, and that pseudo-science may happen to stumble on the truth.³³⁶

³³⁵ Many of the arguments of these authors are based secondary sources and collections of readings in the philosophy of science, rather than the detailed study of primary sources.

³³⁶ Karl Popper, *Conjectures and Refutations* (New York, NY: Routledge, 2010), p. 44.

After World War I, Popper was concerned by the growing popularity of three theories that he believed had more in common with pseudo-science than science: Marxism, Sigmund Freud's psychoanalysis, and Alfred Adler's individual psychology.³³⁷ From his perspective, these theories were constructed and defended in such a way as to sidestep critical evaluation, relying on opaque processes of interpretation and post hoc explanation that ensured that they were never assessed as failures whenever expectations and observations did not align.

In order to confront these defenses, Popper argued that scientists should take risks by making predictions that could be proven false by observation or experiment. He believed that confirming observations, or verification, could never provide proof that a theory was correct, but did allow it to survive and remain in the realm of those that continued to merit consideration. Thus, Popper argued that scientists should make risky assertions that could be clearly interpreted and tested against empirical evidence. He also argued that scientists should avoid making *ad hoc* modifications to their theories, in an effort to avoid refutation based on new interpretations of the theory or the assertion of the presence of a special case.

One can sum up all this by saying that *the criterion of the scientific status of a theory is its falsifiability, or refutability, or testability.*³³⁸

According to Popper, testing conjectures required two intellectual commitments on the part of scientists. First, the empirical world, defined by observations and

³³⁷ Karl Popper, *Conjectures and Refutations* (New York, NY: Routledge, 2010), pp. 44-45.

³³⁸ Karl Popper, *Conjectures and Refutations* (New York, NY: Routledge, 2010), p. 48.

experimental results, trumped theoretical claims. Thus in a clash between evidence and inference, science should demur to the evidence.

...science is distinguished from pseudo-science—or from “metaphysics”—by its *empirical method*, which is essentially *inductive*, proceeding from observation or experiment.³³⁹

Second, because scientists should give priority to evidence, they should not hold onto or defend theories whose predictions have been falsified. As a result, the development of auxiliary hypotheses, search for confounding or intervening variables, or appeals to special cases diminished a theory’s scientific claims, and the theorist’s claim to be acting scientifically.³⁴⁰

By appealing to empiricism and induction, Popper’s demarcation criterion demanded persistent skepticism and constraint on scientists, since there is no *a priori* way of knowing when a conjecture is accurate or at risk of being refuted by the next observation or experiment: Uncertainty about the status of one’s theories was inescapable.

[Popper] claimed that all testing in science has the form of attempting to refute theories by means of observation. And crucially, for Popper it is never possible to confirm or establish a theory by showing its agreement with observations. *Confirmation is a myth*. The only thing an observational test can do is to show that a theory is false. So the truth of a scientific theory can never be supported by observational evidence, not even a little bit, and not even if the theory makes a huge number of predictions that all come out as expected.³⁴¹

³³⁹ Karl Popper, *Conjectures and Refutations* (New York, NY: Routledge, 2010), p. 44.

³⁴⁰ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), p. 61.

³⁴¹ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 58-59.

Popper's ideas have been criticized on a variety of grounds, and alternative demarcation criterion put forward by other philosophers are discussed below. These philosophers found the strict logical structure advanced by Popper to be unnecessarily rigid, too dependent on particular examples from physics, and inconsistent with the history of science.

One of most important criticisms of Popper's demarcation criterion is its emphasis on the beliefs and actions of individual scientists rather than the practice of scientific institutions and communities.

For Popper, a good or great scientist is someone who combines two features, one corresponding to each stage of the cycle. The first feature is an ability to come up with imaginative, creative, and risky ideas. The second is a hard-headed willingness to subject these imaginative ideas to rigorous critical testing. A good scientist has a creative, almost artistic, streak and a tough-minded, no-nonsense streak. Imagine a hard-headed cowboy out on the range, with a Stradivarius violin in his saddlebags. (Perhaps at this point you can see some of the reasons for Popper's popularity among scientists.)...

This raises an interesting question. Empiricist philosophies stress the virtues of open-mindedness, and Popper's view is no exception. But perhaps an open-minded community can be built out of a collection of rather closed-minded individuals. If actual scientists are wedded to their own conjectures, but each is wedded to a different conjecture and would like to prove the others wrong, shouldn't the overall process of conjecture and refutation work? What is wrong with the situation where B's role is to critically test A's ideas? So long as the testing occurs, what does it matter whether A or B does it? One problem is that if everyone is so closed-minded, the results of the test might have no impact on what people believe. Perhaps the young and tender minds of incoming graduate students could be the community's source of flexibility; unsuccessful theories will attract no new recruits and will die with their originators. This would be a rather slow way for science to change (but many would argue that we do see cases like this).³⁴²

³⁴² Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), p. 62.

The other philosophers discussed below have all viewed science as a socially embedded practice and gave greater weight to the behaviors of groups, allowing for the division of labor and specialized roles within scientific communities transforming the pursuit of knowledge as a complex endeavor. As a result, they provide important insights into the design of the intelligence community, and alter the ways in which analysis may be considered scientific based on how groups collectively assess problems, rather than their individual beliefs.

John Ziman

Ziman argued that science was distinct from other sources of knowledge because of its public and social character, rather than its underlying logical structure of argumentation. He noted that the logic used in scientific argument was no different than the logic employed in any other careful discussion used by people in other activities.

In my own experience, one more often detects elementary non sequiturs in the verbal reasoning than actual mathematical mistakes in the calculations that accompany them. This is not said to disparage the intellectual powers of scientists; I mean simply that the reasoning used in scientific papers is not very different from what should use in an everyday careful discussion of an everyday problem.³⁴³

Therefore, rather than demarcate science from pseudo-science based on the logic of refutation, Ziman believed that science was distinct because of the social and psychological aspects of its practice, and that Popper's falsification criterion was too limited to apply to all scientific disciplines.

³⁴³ John Ziman, "What is Science?" in E. D. Klemke, Robert Hollinger, and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), p. 50.

In some disciplines [Karl Popper's falsification criterion] is a good rule of thumb. In chemistry, for example, one should never entertain a hypothesis that could not in principle, by some means or another, be shown to be quite wrong. But in historical disciplines such as palaeontology this is too restrictive, and in cosmology it would be completely frustrating. Again, the philosophy of science has been so focused on the physical sciences that it makes too little allowance for the innate variability of biological organisms, which limits the possibilities of meaningful quantification.³⁴⁴

Ziman's alternative view of science focused on how knowledge was developed, advanced, and argued about in a transparent fashion among peers. This notion of public knowledge extended beyond the act of publication and extended into the communal nature of science itself. In Ziman's view, scientists engaged in an intellectual process that was simultaneously competitive and cooperative. Competition motivated individuals to jealously watch one another's work and challenge their conclusions, resulting in a collective effort to evolve new knowledge by generating novel theories and eliminating ineffective ones. This competition at the microlevel created cooperation at the macrolevel, as communities of scientists moved through the chain of logical deduction, to empirical observation, to inductive inference, and back again.

The defect of the conventional philosophical approach to science is that it considers only two terms in the equation. The scientist is seen as an individual, pursuing a somewhat one-sided dialogue with taciturn nature. *He* observes phenomena, notices regularities, arrives at generalizations, deduces consequences, et cetera and eventually, Hey Presto! A law of nature springs into being. But it is not like that at all. The scientific enterprise is corporate. It is not merely, in Newton's incomparable phrase, that one stands on the shoulders of giants, and hence can see a little farther. Every scientist sees through his own eyes—and also through the eyes of his predecessors and colleagues. It is never one individual that goes through all the steps in the logico-inductive chain; it is a group of individuals, dividing their labor but continuously and jealously checking

³⁴⁴ John Ziman, "Getting Scientists to Think About What They are Doing," *Science and Engineering Ethics*, Vol. 7, No. 2 (2001), p. 168.

each other's contributions. The cliché of scientific prose betrays itself "Hence we arrive at the conclusion that..." The audience to which scientific publications are addressed is not passive; by its cheering or booing, its bouquets or brickbats, it actively controls the substance of the communications it receives.

In other words, scientific research is a social activity.³⁴⁵

By emphasizing the structure and practices of scientific communities, Ziman focused on social relations and the practices of peer groups. In doing so, practices that have traditionally been attributed to non-scientific ventures, such as relying on apprenticeships, creativity, and the learning of identities, are reimagined as integral parts of the scientific communities and established social activities.

The fact is that scientific investigation, as distinct from the theoretical content of any given branch of science, is a practical art. It is not learnt out of books, but by imitation and experience....

To understand the nature of science, we must look at the way in which scientists behave towards one another, how they are organized, and how information passes between them. The young scientist does not study formal logic, but he learns by imitation and experience a number of conventions that embody strong social relationships. In the language of sociology, he learns to play his *role* in a system by which knowledge is acquired, sifted, and eventually made public property.³⁴⁶

Thus, Ziman shifted the criterion for demarcation away from Popper's logic of empiricism and falsification, to a complex mixture of individual intellect, psychology, and sociology that discursively generated knowledge through open cooperation and competition. The importance of this shifting emphasis on the social structure of science was described by Godfrey-Smith as allowing scientific communities to develop in two

³⁴⁵ John Ziman, "What is Science?" in E. D. Klemke, Robert Hollinger, and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), p. 51.

³⁴⁶ John Ziman, "What is Science?" in E. D. Klemke, Robert Hollinger, and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), pp. 50, 51-52.

dimensions: The first dimension coordinated the activities of scientists working contemporaneously while the second dimension ensured the intergenerational transfer of knowledge that enables its accumulation over time.

We need to focus on the development and structure of a *socially organized way of carrying out* the basic scientific strategy.

The distinctive features of science as a social structure are found along two different dimensions. One has to do with the organization of work at a given time. Here we find the suggestion that science has developed a reward system and an internal culture that generate an efficient mixture of competition and cooperation, and a beneficial division of scientific labor across different approaches to a problem.... The general argument is that science (construed narrowly, as involving a particular social structure) is able to *coordinate the energies* of diverse individuals in an effective way.

The other dimension has to do with the relationships between different times, and with the transmission of ideas between scientific generations. The crucial feature we find along this dimension is that scientific work is cumulative. Each generation builds on the work of predecessors; current workers “stand on the shoulders” of earlier workers, as Isaac Newton once put it. This requires both trustworthy ways of transmitting ideas across time and (again) a reward system that makes it worthwhile to carry on where earlier workers left off. With a social structure of this kind, the “dialogue between the imaginative and the critical voices” can become a real dialogue. We have social mechanisms in place that reliably bring about the checking and scrutinizing of ideas. To use a phrase suggested to me by Kim Sterelny, we get an “engine of self-correction” to accompany the speculative side of scientific thinking. In a situation like this, we can have a true division of labor in how the basic empiricist pattern is manifested. Some dogmatic and bloody-minded individuals can work within the system and even play a potentially useful role, provided that flexibility and open-mindedness is found in the community as a whole.³⁴⁷

Thomas Kuhn

Kuhn also criticized Popper’s view of demarcation on the grounds that it failed to account for actual scientific practice, particularly the extent to which scientists continue

³⁴⁷ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 224-225.

to believe in, and employ, falsified theories. Kuhn challenged whether science was a purely rational, logical, and impersonal pursuit as Popper had suggested. Instead, he advanced a version of science that sought to solve problems, or more specifically, find solutions to puzzles based on the tension between two competing processes that he referred to as normal and extraordinary (or revolutionary) science. When taken together, these two modes offered an evolutionary model of science that shifted between exploiting existing theories by expanding the number of puzzles they could solve, and exploring alternative theories upon encountering puzzles that existing theories could not solve.³⁴⁸

Kuhn argued that under the conditions of normal science, researchers sought to solve puzzles by applying established theories, and, importantly, if they failed, the skill of the scientist and not the validity of the theory was questioned.

...when engaged with a normal research problem, the scientist must *premise* current theory as the rules of his game. His object is to solve a puzzle, preferably one at which others have failed, and current theory is required to define that puzzle to guarantee that, given sufficient brilliance, it can be solved. Of course the practitioner of such an enterprise must often test the conjectural puzzle solution that his ingenuity suggests. But only his personal conjecture is tested. If it fails the test, only his own ability not the corpus of current science is impugned. In short, though tests occur frequently in normal science, these tests are of a peculiar sort, for in the final analysis it is the individual scientist rather than current theory which is tested.³⁴⁹

From Kuhn's perspective, the goal of normal science was to expand the number of puzzles that a theory solved, extending its claim to provide fundamental insights based on increasing generalization. In their efforts to extend the reach and power of a theory,

³⁴⁸ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 75-101.

³⁴⁹ Thomas S. Kuhn, "Logic of Discovery or Psychology of Research?" in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton & Company, 1998), p. 12.

Kuhn argued that scientists often went to extraordinary lengths to reconstruct problems in such a way as to make the theory correct rather than abandon it in favor of alternatives as Popper advocated. Godfrey-Smith noted how the process of puzzle solving occurring within normal science ironically produced the germs of extraordinary science.

The normal scientist tries to use the tools and concepts provided by the paradigm to describe, model, or create new phenomena. The “puzzle” is trying to get a new case to fit smoothly into the framework provided by the paradigm. Kuhn used the term “puzzle” rather than “problem” for a reason. A puzzle is something we have not yet solved but which we think does have a solution. A problem might, for all we know, have no solution. Normal science tries to apply the concepts provided by a paradigm to issues that the paradigm suggests *should* be soluble....

A normal scientist does, Kuhn thinks, spend a lot of time on topics that look insignificant from the outside.... But it is this close attention to detail—which only the well-organized machine of normal science makes possible—that is able to reveal *deep* new facts about the world. I think Kuhn felt a kind of awe at the ability of normal science to home in on topics and phenomena that look insignificant from outside but which turn out eventually to have huge importance. And although the normal scientist is not *trying* to find phenomena that lead to paradigm change—far from it!—these detailed discoveries often contain the seeds of large-scale change and the destruction of the paradigm that produced them.³⁵⁰

Kuhn balanced his view of normal science as an essentially conservative enterprise with a radical mode of scientific research: extraordinary science. When the number of puzzles that a theory could not solve accumulated and alternative ideas became available, extraordinary science occurred. During these periods, rival theories merited investigation, and the paradigms within which normal science operated were reevaluated and often abandoned. Thus, extraordinary science changed how scientists

³⁵⁰ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 81-82.

defined problems, characterized puzzles and their solutions, and pursued their research by altering their research methods, assumptions, and treatment of data.³⁵¹

In a scientific revolution, as in a political one, rules break down and have to be rebuilt afresh. If you look at two pieces of scientific work *across* a revolutionary divide, it will not be clear whether there has been progress from earlier to later. It might not even be clear how to compare the theories or pieces of work *at all*—they may look like fundamentally different kinds of intellectual activity. The people on different sides of the divide will be “speaking different languages.” In the climax of his book, Kuhn says that workers in different paradigms are living in different worlds.³⁵²

Kuhn believed that the tension between normal and extraordinary science served as the engine for discovery and progress within science, and that Popper’s logical demarcation criterion failed to capture this dynamic by focusing exclusively on rare cases of scientific revolutions.³⁵³

...the tests which [Popper] emphasizes are those which were performed to explore the limitations of accepted theory or to subject a current theory to maximum strain. Among his favourite examples, all of them startling and destructive in their outcome are Lavoisier’s experiments on calcination, the eclipse expedition of 1919, and the recent experiments on parity conservation. All, of course, are classic tests, but in using them to characterize scientific activity [Popper] misses something terribly important about them. Episodes like these are very rare in the development of science. When they occur, they are generally called forth either by a prior crisis in the relevant field...or by the existence of a theory which competes with the existing canons of research (Einstein’s general relativity). These are, however, aspects of or occasions for what I have elsewhere called “extraordinary research,” an enterprise in which scientists do display very many of the characteristics [Popper] emphasizes,

³⁵¹ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 75-79.

³⁵² Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), p. 88.

³⁵³ Thomas S. Kuhn, “Logic of Discovery or Psychology of Research?” in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton & Company, 1998), pp. 12-13.

but one which, at least in the past, has arisen only intermittently and under quite special circumstances in any scientific specialty.³⁵⁴

As a result, Kuhn argued that Popper missed the essential social underpinnings of normal science and the paradigms they established and promoted.

Critics of Kuhn suggest that he saw the majority of science as unimaginative, technical, and irrationally defensive, but viewing science as an evolutionary process suggests otherwise.³⁵⁵ Instead, the protection, promotion, and proliferation of a theory through training and socialization allowed for the full development of ideas that would otherwise experience premature deaths if discarded at the first sign of falsification.³⁵⁶ Moreover, in order to solve puzzles, a common framework for determining what they were and what solutions looked like was required—the exact features that normal science provided.

No puzzle-solving enterprise can exist unless its practitioners share criteria which, for that group and for that time, determine when a particular puzzle has been solved. The same criteria necessarily determine failure to achieve a solution, and anyone who chooses may view that failure as the theory to pass a test.³⁵⁷

Indeed, Kuhn's notion of scientific paradigms played an important role in consolidating knowledge and approaches to science that allow for practitioners to extend the reach of their theories. By sharing fundamental assumptions, language, methodological standards, and approaches to research design, researchers within a field

³⁵⁴ Thomas S. Kuhn, "Logic of Discovery or Psychology of Research?" in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton & Company, 1998), pp. 12-13.

³⁵⁵ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 79-86.

³⁵⁶ Isaac Ben-Israel, "Philosophy and Methodology of Intelligence: The Logic of Estimate Process" *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 676-678.

³⁵⁷ Thomas S. Kuhn, "Logic of Discovery or Psychology of Research?" in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton & Company, 1998), p. 15.

can make progress by sharing each other's work. As Godfrey-Smith noted, paradigms consolidated knowledge and experience, allowing science to progress by providing a foundation upon which researchers could build.

A paradigm's role is to *organize* scientific work; the paradigm coordinates the work of individuals into an efficient collective enterprise. A key feature that distinguishes normal science from other kinds of science for Kuhn is the absence of debate about fundamentals. Because scientists doing normal science agree on these fundamentals, they do not waste their time arguing about the most basic issues that arise in their field. Once biologists agree that genes are made of DNA, they can focus and coordinate their work on how specific genes affect the characteristics of plants and animals. Once chemists agree that understanding chemical bonding is understanding the interactions between the outer layers of electrons within different atoms, they can work together to investigate when and how particular reactions will occur. Kuhn places great emphasis on this "consensus-forging" role of paradigms. He argues that without it, there is no chance for scientists to achieve a really detailed and deep understanding of phenomena. Detailed work and revealing discoveries require cooperation and consensus. Cooperation and consensus require *closing off debate about fundamentals*.³⁵⁸

Kuhn's criterion for demarcation rested on how scientists, as a community, identified and solved puzzles through a tension-filled, evolutionary process. While this process may have been irrational at the individual level, in contradiction to Popper's emphasis on the rational behavior of individuals, it resulted in a rational, progressive development of knowledge punctuated by dramatic revolutions as paradigms replaced one another. Kuhn's emphasis on dynamics resulted from the competition between normal and exceptional science, and allowed for science to proceed in a rational order without the need for individual scientists to be rational individuals.

³⁵⁸ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 80-81.

First, in some ways Kuhn's view of science has an "invisible hand" structure. The Scottish political and economic theorist Adam Smith argued in the *Wealth of Nations*...that individual selfishness in economic behavior leads to good outcomes for society as a whole. The market is an efficient distributor of goods to everyone, even though the people involved are each just out for themselves. Here we have an apparent mismatch between individual-level characteristics and the characteristics of the whole; selfishness at one level leads to the general benefit. But the mismatch is only apparent; it disappears when we look at the consequences of having a large number of individuals interacting together. We see something similar in Kuhn's theory of science: narrow-mindedness and dogmatism at the level of the individual lead to intellectual openness at the level of science as a whole. Anomaly and crisis produce such stresses in the normal scientist that an especially wholesale openness to novelty is found in revolutions....

A biological analogy can also be found in the case of Kuhn. During the 1970s, the biologists Stephen Jay Gould, Niles Eldredge, and others argued that the large-scale pattern seen in much biological evolution is "punctuated equilibrium".... A lineage of organisms in evolutionary time will usually exhibit long periods of relative stasis, in which we see low-level tinkering but little change to fundamental structures. These periods of stasis or equilibrium are punctuated by occasional periods of much more rapid change in which new fundamental structures arise. (Note that "rapid" here means taking place in thousands of years rather than millions.) The rapid periods of change are disorderly and unpredictable when compared to the simplest kind of natural selection in large populations. The periods of stasis also feature a kind of "homeostasis" in which the genetic system in the population tends to *resist* substantial change.

The analogy with Kuhn's theory of science is striking. We have the same long periods of stability and resistance to change, punctuated by unpredictable, rapid change to fundamentals.³⁵⁹

Irme Lakatos

Lakatos also challenged Popper's emphasis on falsification, but differed from Ziman's argument about public knowledge and Kuhn's belief about the tensions between normal and extraordinary science. Lakatos questioned the viability of testing theories piecemeal, significantly undercutting the practicality of Popper's strict appeals to

³⁵⁹ Peter Godfrey-Smith *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 99-100.

observational experience by casting doubt on the relationship between empirical facts and theoretical predictions. Instead, he argued that empirical observations and experiments tested the claims of research programs, which consisted of stable cores and dynamic, auxiliary hypotheses.

Lakatos argued that Popper had undermined conventional notions of induction based on his falsification criteria—after all, if the collection of facts could not prove a theory true, what then was the relationship between facts and theory? He noted that the ironic, logical consequence of Popper’s emphasis on falsification actually disconnected theory and evidence in an unexpected way by transforming substantive questions about specific phenomenon into general methodological or epistemological ones. As a result, Popper’s demarcation criterion did not distinguish scientific ideas from pseudo-scientific ideas, but scientific methods from non-scientific methods.

...in 1934, Karl Popper, one of the most influential philosophers of our time, argued that the mathematical probability of all theories, scientific or pseudoscientific, given any amount of evidence is zero. If Popper is right, scientific theories are not only equally unprovable but also equally improbable. A new demarcation criterion was needed and Popper proposed a rather stunning one. A theory may be scientific even if there is not a shred of evidence in its favour, and it may be pseudoscientific even if all the available evidence in its favour. That is, the scientific or non-scientific character of a theory can be determined independently of the facts. A theory is “scientific” if one is prepared to specify in advance a crucial experiment (or observation) which can falsify it, and it is pseudoscientific if one refuses to specify such a “potential falsifier.” But if so, we do not demarcate scientific from pseudoscientific ones, but rather scientific method from non-scientific method. Marxism, for a Popperian, is scientific if the Marxists are prepared to specify the facts which, if observed, make them give up Marxism. If they refuse to do so, Marxism becomes pseudoscience.³⁶⁰

³⁶⁰ Imre Lakatos, “Science and Pseudoscience,” in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton and Company, 1998), pp. 22-23.

Lakatos credited Popper for recognizing that conformation with facts was insufficient criteria for accepting theories, and that the boundary between science and pseudo-science lied in the ways in which ideas were expressed. Yet, he also agreed with Kuhn on the fact that scientists rarely abandoned their theories based on falsification alone.

Scientists have thick skins. They do not abandon a theory merely because facts contradict it. They normally either invent some rescue hypothesis to explain what they then call a mere anomaly or, if they cannot explain the anomaly, they ignore it, and direct their attention to other problems. Note that scientists talk about anomalies, recalcitrant instances, not refutations. History of science, of course, is full of accounts of how crucial experiments allegedly killed theories. But such accounts are fabricated long after the theory had been abandoned. Had Popper ever asked a Newtonian scientist under what experimental conditions he would abandon Newtonian theory, some Newtonian scientists would have been exactly as nonplussed as are some Marxists.³⁶¹

And while Lakatos agreed with Kuhn on the durability of theories, he was particularly troubled by Kuhn's view that definitions of science were socially constructed and embedded within particular communities and paradigms, which did not allow for a universal, objective standard of demarcation.³⁶² Indeed, Lakatos was so interested in restoring the status of science as a rational pursuit that other philosophers have been puzzled by his treatment of history.

He saw Kuhn's influence as *destructive*—destructive of reason and ultimately dangerous to society. For Lakatos, Kuhn had presented scientific change as a fundamentally irrational process, a matter of “mob psychology,”... a process where the loudest, most energetic, and most numerous voices would prevail regardless of reasons.... Lakatos... saw

³⁶¹ Imre Lakatos, “Science and Pseudoscience,” in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton and Company, 1998), p. 23.

³⁶² Imre Lakatos, “Science and Pseudoscience,” in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton and Company, 1998), p. 23.

the disorder in Kuhn's picture as no more than dangerous chaos. But Lakatos also saw the force of Kuhn's historical arguments. So his project was to rescue the rationality of science from the damage Kuhn had done....

Lakatos argued that historical case studies should be used to assess philosophical views of science. Fine, so far. But he also said that we should write "rational reconstructions" of the historical episodes, in which scientists' decisions are made to look as rational as possible. We should then separately (or in footnotes) point out places where the rational reconstruction is not an accurate description of what actually went on. So it is OK to *deliberately misrepresent* what happened in the past, so long as the footnotes set things straight.... What matters most is that in the main discussion we are able to spin a story in which the scientific decisions come out looking rational.³⁶³

In order to resolve his differences between Popper and Kuhn, Lakatos argued that science proceeded by developing and testing entire research programs, not individual theories. While research programs were similar to Kuhn's paradigms, there was one crucial difference, i.e. the number operating in scientific fields at any moment in time. Kuhn envisioned scientific disciplines to be adherents to one paradigm at time—thus when one replaced another, a revolution occurred. By contrast, Lakatos believed that scientific disciplines contained populations of competing paradigms operating at all times, allowing them to serve a rationalizing role over the field.³⁶⁴

Research programs were composed of a core set of theories surrounded by auxiliary hypotheses that protected the core from refutation. Research programs also provided heuristics that guided the activities of scientists, by supplying them with schemes for problem solving, the handling of anomalies, and even converting contradictory facts into supporting evidence.

³⁶³ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 103-104.

³⁶⁴ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), p. 104.

Science is not simply trial and error, a series of conjectures and refutations. “All swans are white” may be falsified by the discovery of one black swan. But such trivial trial and error does not rank as science. Newtonian science, for instance, is not simply a set of four conjectures—the three laws of mechanics and the law of gravitation. These four laws constitute on the “hard core” of the Newtonian programme. But this hard core is tenaciously protected from refutation by a vast “protective belt” of auxiliary hypotheses. And, even more importantly, the research programme also has a “heuristic,” that is, a powerful problem-solving machinery, which, with the help of sophisticated mathematical techniques, digests anomalies and even turns them into positive evidence.³⁶⁵

According to Lakatos, demarcation between science and pseudo-science rested on the ability to evaluate and challenge entire research programs, not the individual theories or auxiliary hypotheses that protected them. Research programs that prove capable of predicting novel, unexpected, or surprising results—either in the form of new, unforeseen discoveries or facts that rival explanations fail to predict—are progressive and scientific. Alternatively, research programs that invent new auxiliary hypotheses to insulate their core from falsification, and are consistently reacting to the discovery of new facts discovered by rival programs, should be regarded as regressive and drifting toward pseudo-science.³⁶⁶

Paul Thagard

Thagard continued the theme advanced by Ziman and Kuhn by employing social criteria and practices as the basis of demarcation. He argued that determining a theory’s scientific status was found in the social practices of its adherents as they changed over time.

³⁶⁵ Imre Lakatos, “Science and Pseudoscience,” in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton & Company, 1998), pp. 23-24.

³⁶⁶ Imre Lakatos, “Science and Pseudoscience,” in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton & Company, 1998), pp. 25-26.

In developing his demarcation criteria, Thagard considered the scientific status of astrology and arguments against its inclusion in science in order to challenge popular definitions of science while providing his own. Thagard dismissed several criticisms of astrology's scientific status by noting that they rested on arguments that were not relevant to the problem of demarcation. For example, Thagard argued that the historical origins of astrology being grounded in religion and myth were not relevant. Neither was the lack of a theory that would explain the results of statistical tests on the correspondence between astrological predictions and outcomes, nor were the psychological motivations of its adherents relevant.³⁶⁷

These objections do not show that astrology is a pseudoscience. First, origins are irrelevant to scientific status. The alchemical origins of chemistry and the occult beginnings of medicine are as magical as those of astrology, and historians have detected mystical influences in the work of many great scientists, including Newton and Einstein. Hence astrology cannot be condemned simply for the magical origins of its principles. Similarly, the psychology of popular belief is also in itself irrelevant to the status of astrology: people often believe even good theories for illegitimate reasons, and even if most people believe astrology for personal, irrational reasons, good reasons may be available. Finally the lack of a physical foundation hardly marks a theory as unscientific. Examples: when Wegener proposed continental drift, no mechanism was known, and a link between smoking and cancer has been established statistically though the details of carcinogenesis remain to be discovered.³⁶⁸

Thagard also noted that resorting to empirical observation and claims of success or failure were not sufficient for demarcating science from pseudo-science. Instead, he offered three criteria as necessary. First, a theory must be capable of giving structure to a

³⁶⁷ Paul R. Thagard, "Why Astrology is a Pseudoscience," in E. D. Klemke, Robert Hollinger, and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), pp. 68-69.

³⁶⁸ Paul R. Thagard, "Why Astrology is a Pseudoscience," in E. D. Klemke, Robert Hollinger, and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), p. 68.

problem through the identification of concepts and their measurement. Additionally, a community of adherents must generally agree upon the theory's principles and applications to problems. This community must share a concern over the occurrence of anomalies between the theory's predictions and observations, actively seek to challenge the theory with new observations and evidence, and seek to extend the range of problems or facts explained by their theory. Finally, if the community exists in an environment where rival theories are available, the community must compare their relative strengths and capabilities.³⁶⁹ Thus Thagard concluded that the demarcation criteria for theories are historical and dynamic.

A theory or discipline which purports to be scientific is pseudoscientific if and only if:

- (1) It has been less progressive than alternative theories over a long period of time, and faces many unsolved problems; but
- (2) the community of practitioners makes little attempt to develop the theory towards solutions of the problems, shows no concern for attempts to evaluate the theory in relation to others, and is selective in considering confirmations and disconfirmations.³⁷⁰

Determining a theory's scientific status results from observing how its adherents behave as communities. An isolated group possessing one set of beliefs with no ability to gather relevant data or conduct experiments may be regarded as scientific if they remain curious, open, and challenge their beliefs when new opportunities to do so arrive.

Alternatively, a group holding the same beliefs, but embedded in a diverse community of

³⁶⁹ Paul R. Thagard, "Why Astrology is a Pseudoscience," in E. D. Klemke, Robert Hollinger, and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), p. 73.

³⁷⁰ Paul R. Thagard, "Why Astrology is a Pseudoscience," in E. D. Klemke, Robert Hollinger, and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), pp. 70-71.

ideas and observational opportunities, but actively distancing themselves from challenges, may be considered pseudoscientific.

Demarcation and Intelligence

Each of the preceding perspectives on demarcation provides insights into intelligence analysis, and suggests how science, in both its ideal and actual forms, may serve as a model for the intelligence community. What follows is a provisional description of analytic tradecraft and organization based on the models of science provided by Popper, Ziman, Kuhn, Lakatos, and Thagard.

Popper's emphasis on falsification has been well integrated into analytic tradecraft, particularly in the form of Alternative Competing Hypotheses (ACH) and other Structured Analytic Techniques (SATs) that see intelligence analysis as a form of hypothesis generation and testing.³⁷¹ These approaches encourage analysts to take risks by constructing intelligence assessments as testable, falsifiable conjectures to the maximum extent possible, and encourage analysts to abandon or downgrade the likelihood of hypotheses whenever collection provides disconfirming evidence.

Popper's emphasis on falsification also suggests the need for close working relationships between analysts and collectors. In this relationship, collection would seek to falsify analytic judgments and invalidate assumptions in order to test the strength of their claims. As Ben-Israel argued:

³⁷¹ Isaac Ben-Israel, "Philosophy and Methodology of Intelligence: The Logic of Estimate Process," *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 708-711; Richards J. Heuer, Jr., *Psychology of Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 1999), pp. 95-110; Morgan D. Jones, *The Thinker's Toolkit: 14 Powerful Techniques for Problem Solving* (New York, NY: Crown, 1998), pp. 178-216; and conversation with James Bruce, RAND, August 13, 2012.

The intelligence researcher, like the scientist, must begin his research by *formulating the problem* which troubles him. The first stage is to create a list of hypotheses as possible solutions. Then the possible solutions must be checked, one by one, by searching for falsifying information... .

Information-gathering should thus be directed according to the relevant research problem. Information should not be gathered “blindly.” Instead, observation reports which might falsify the listed hypotheses should be sought. If a hypothesis is refuted (and this is possible only after the refuting information itself is checked), we must erase it from our list and go to the next. The convergence toward a “solution” is, therefore, a process of elimination....

It is not denied that reality is sometimes richer than imagination, and that its study can open our eyes to unimagined possibilities. This is not, however, the main aim of information-gathering. Information is gathered to help refute false hypotheses. It is always guided by the problem to be solved, and is never blind.³⁷²

Thus, rather than have analysts sift through volumes of information in search of key pieces of data that fit into their analytic frameworks, reaffirming the effects of confirmation biases, or have analysts serve as vehicles for marketing collection capabilities to consumers, analysts would aid collectors in focusing their activities and use of resources against very specific, tailored requirements specifically selected to test analytic assessments stated as hypotheses.³⁷³

Ziman’s perspective on science as a communal, social activity can be seen in the design of the intelligence community and its analytic production processes. The structure of the US intelligence community is characterized by semi-autonomous agencies and offices that simultaneously develop independent, competing assessments of intelligence

³⁷² Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 678-679.

³⁷³ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 678-679; William M. Nolte, “Intelligence Analysis in an Uncertain Environment,” in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), p. 412; and Josh Kerbel and Anthony Olcott, “Synthesizing with Clients, Not Analyzing for Customers,” *Studies in Intelligence*, Vol. 54, No. 4 (December 2010), pp. 17-18.

problems and coordinated, communal perspectives through the drafting of NIEs, National Intelligence Assessments (NIAs), products developed by fusion centers, and more.³⁷⁴ In fact, Bruce noted that the coordination process constituted the closest communal aspect of scientific practice within intelligence analysis, but is often overwhelmed by social and organizational factors that diminish its effectiveness with respect to identifying and correcting analytic errors.

Coordination is the only explicit step in the analytical process that already provides potential self-corrective mechanisms. But the coordination process, regrettably, is too rarely used for this purpose. For the most part, interagency and intra-agency coordination have been corrupted into a linguistic exercise. Analysts and managers seek agreeable prose, words that may (or may not) help a policymaker but are crafted to get agreement among the parties in order to publish them. The primacy of finding just the right *words* to facilitate going to press necessarily subverts an otherwise invaluable epistemological process.³⁷⁵

In spite of these difficulties, however, the intelligence community possesses the capacity to replicate the underlying social structure of science as a two-level process characterized by competition and cooperation—a point that has become increasingly apparent with the development of a variety of SATs, particularly Devil’s Advocacy and Team B studies.³⁷⁶

³⁷⁴ Thomas Fingar, *Reducing Uncertainty: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011), p. 60.

³⁷⁵ James B. Bruce, “Making Analysis More Reliable: Why Epistemology Matters to Intelligence,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 185.

³⁷⁶ Roger Z. George, “Fixing the Problem of analytic Mind-Sets: Alternative Analysis,” *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), pp. 385-404; Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York, NY: Columbia University Press, 2007), pp. 19-52; Central Intelligence Agency, *A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis* (Washington, DC: Central Intelligence Agency, March 2009); Richards J. Heuer, Jr., and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly, 2011).

Kuhn's views of science are less well integrated into analytic tradecraft than Popper's, but highly relevant for performing and evaluating analysis and thinking about intelligence failures. Viewing intelligence analysis through Kuhn's framework suggests that there are two modes in which analysts operate. The first replicates Kuhn's normal science and might be called normal analysis. Normal analysis operates by developing and promoting analytic paradigms that include frameworks for defining and bounding analytic problems, strategies for collection, key assumptions, and theoretical concepts and means of measurement, e.g. political stability and legitimacy or military readiness. The use of these frameworks encourages consistency in reporting and assessments, and enables the development and evaluation of expertise within peer groups. Alternatively, the second mode of intelligence analysis, extraordinary analysis, based on Kuhn's extraordinary science, occurs when the normally employed analytic frameworks or paradigms require revision. Under extraordinary analysis, consistency with prior assessments declines and alternative ideas and approaches receive greater attention in order to explain the accumulation of anomalies in collected data.

The tension between normal and extraordinary modes of intelligence analysis characterizes contemporary perspectives on intelligence failures, particularly with respect to individual mindsets and group dynamics.³⁷⁷ Members of the intelligence community

³⁷⁷ Roger Z. George, "Fixing the Problem of Analytic Mind-Sets: Alternative Analysis," *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), pp. 385-404; Jack Davis, "Why Bad Things Happen to Good Analysts," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 157-170; Susan G. Straus, Andrew M. Parker, James B. Bruce, and Jacob W. Dembosky, *The Group Matters: A Review of the Effects of Group Interaction on Processes and Outcomes in Analytic Teams* (Santa Monica, CA: RAND, April 2009), and Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly Press, 2011).

have argued that normal theories may be employed with great success for long periods of time and across many cases, only to fail without warning. As Jack Davis noted:

...there is no known theory, practice, or methodological tool for infallible determination of whether a normal or exceptional course of events lies ahead.³⁷⁸

In keeping with Kuhn's concepts, scientific revolutions resulting from changes between normal to extraordinary science are characterized by great discoveries and paradigm shifts. However, in the intelligence community, these discoveries may be anxious moments of vulnerability, failure, and strategic surprise—occurring upon the realization that the accepted analytic frameworks that define intelligence problems, key assumptions, and data sources have led analysts and consumers to develop an inaccurate and deeply flawed understanding of a problem.

Finally, Kuhn's suggestion that science is not a rational process, or at least that scientists may not be individually rational, carries significant weight regarding the structure and management of the entire intelligence community. Many of the institutional challenges facing intelligence in the management of producer/consumer relations come from the need to maintain objectivity and policy neutrality. On this point, the dispassionate pursuit of knowledge and the ability to hedge against bias in science are seen as points worthy of emulation within the intelligence community.³⁷⁹ Yet, this is also a mischaracterization of science—one that confuses the role and relations of individual

³⁷⁸ Jack Davis, "Why Bad Things Happen to Good Analysts," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), p. 163. Also see Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York, NY: Columbia University Press, 2007), pp. 53-65.

³⁷⁹ Jennifer E. Sims, "Decision Advantage and the Nature of Intelligence Analysis," in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), p. 390.

scientists and the scientific community itself. Because Kuhn, and others, described how individual scientists may not be rational or dispassionate, but personally and professionally invested in particular research approaches or conclusions, it is the institutional and procedural structures of science that do the heavy lifting of error correcting, fact checking, and depersonalization of arguments between competing theories. As Ben-Israel noted:

Scientific objectivity is not based on the objectivity of [the] individual scientist, but only on that of *scientific establishments*—research institutes, scientific books and periodicals, scientific conventions, laboratories and scientific language. Such establishments enable every scientist to carry out his own tests and verify results reported in scientific literature by other researchers.

Nor is there any *a priori* reason for assuming that a natural scientist is more detached from his research (sometimes encompassing a lifetime) than an intelligence researcher. A physicist or engineer, who believes he is about to crack a secret of nature or build a revolutionary machine is no less excited or involved than an intelligence estimator or social scientist. Scientists, exactly like intelligence researchers, are human beings, prejudiced and prone to error. There are scientists who abuse their scientific ability and reputation, and no one takes—or should take—their objectivity for granted. Or, in words of a wise man, we should show “more respect to *science* than to *scientists*.”³⁸⁰

Lakatos’s contributions to intelligence studies are equally insightful for thinking about intelligence failure and competing assessments. Lakatos’s philosophy of science is relevant in two important ways with respect to intelligence analysis. First, his emphasis on the prediction of novel facts and new discoveries, rather than logical falsification, provides a different metric for evaluating intelligence estimates. Instead of evaluating analysts’ judgments based on the weight of contradictory observations, per Popper’s

³⁸⁰ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 687.

recommendation, assessments should be valued based on their correspondence with new collection. While this would leave analysts vulnerable to confirmation bias, it would constantly challenge the completeness of their models and explanations by comparing their expectations with “out-of-sample” data. Thus, while Popper’s application to intelligence favored placing collection in a subservient position to analysis, Lakatos’s emphasis on the prediction of novel facts suggests that collectors should to pursue novel and wildcat collection efforts to test whether unexpected data sources and methods provide evidence that is consistent with analytic theory.

Lakatos’ second contribution is more subtle, but quite powerful with respect to denial, deception, politicization, and producer/consumer relations. His focus on research programs as a core of assumptions and theories surrounded by protective belts of auxiliary hypotheses and *ad hoc* interpretations of data explains the hardening of assessments and their immunity to criticism and falsification. By dividing intelligence estimates, or consumers’ own expectations, into commitments to a core analytic conclusion and the protective measures that defend it from falsification, paths to politicization and the ambiguities created by denial and deception become clearer. Advocates of particular analytic conclusions may argue that “absence of evidence is not evidence of absence” with respect to a target’s intentions or capabilities, transforming the inability to uncover something as proof of its existence.³⁸¹ While the strategic context of international relations and other competitive activities permits arguments constructed on the basis of active and successful manipulation of the collection and analytic processes by

³⁸¹ Gordon R. Mitchell, “Team B Intelligence Coups,” *Quarterly Journal of Speech*, Vol. 92, No. 2 (2006), p. 155.

outside agents, the shift in epistemological structure must be acknowledged and serve as a warning to producers and consumers alike. Specifically, assessments that invoke deception as the basis of their claims have largely sidestepped empirical science by operating outside the bounds of available evidence—an act not without justification, but difficult to support on scientific grounds alone.

While there may be good and often necessary reasons to elevate inferences over evidence in the production of intelligence, doing so should alert producers and consumers to the special weight being carried by auxiliary hypotheses and assumptions and their role in protecting an unverified core. In these cases, one of three difficult truths must be accepted:

- The target is effectively manipulating intelligence sources and methods, explaining why the available evidence should not be believed. This calls into question the effectiveness of intelligence operations, counterintelligence, and the veracity of other assessments based on potentially compromised sources and methods;
- Producers and/or consumers are so overly committed to proving a particular outcome that they have manipulated the evidence to conform to a predetermined conclusion. This is a particularly problematic result because it would indicate that deep-seated, often immutable identity roles are driving

analysis and may not be acceptably challenged within the existing institutional structure,³⁸²

- Both are occurring simultaneously.

Lakatos's research program-based science suggests that if an assessment's advocates dismiss evidence on the grounds that it is manipulated, deceptive, or simply misinterpreted, then their alternative framework should be able to predict new and novel facts resulting from future collection activities, i.e. explain observations that rival theories cannot. If this does not occur, then the defensive, *ad hoc* protection of assumptions and conclusions should be seen as regressive maneuvering in support of a favored outcome, rather than an honest analytic effort that cannot be countered by appeals to evidence and additional collection. When this occurs, evaluations of estimates should examine the organizational independence and operational isolation of analysts, which encourage and enable them to safeguard ideological commitments rather than demand the open-minded treatment of the available evidence and its relationship with collection strategies and capabilities.

Thagard's demarcation criterion elaborates on the research program approach developed by Lakatos. A scientific approach to intelligence analysis, from Thagard's

³⁸² Theories of identity argue that perceptions of information are often manipulated in order to reinforce existing beliefs about in-group and out-group distinctions and comparisons. Thus, if intelligence analysts defined their identity around the identification of threats from foreign actors, they are likely to interpret signals from foreign actors to be threatening, and the more benign the appearance of the signal, the more deceptive and threatening the source. Moreover, if these identity roles and beliefs in in-group and out-group differences achieve "doxic" status, i.e. they are part of the belief system that defines group membership, they cannot be challenged by group members. These aspects of identity roles are discussed in greater depth in Chapter 9. See Peter J. Burke and Jan E. Stets, *Identity Theory* (New York, NY: Oxford University Press, 2009), pp. 28-29; and Steph Lawler, *Identity: Sociological Perspectives* (New York, NY: Policy, 2009), p. 127.

perspective, would be relatively easy to identify. First, it would rest upon the use of theories or analytic frameworks that provide structure to intelligence problems, serving the same role Kuhn's paradigms, or Lakatos's research programs. Intelligence analysts must be capable of applying frameworks consistently to the same problem, and ensure that others who employ the same theories and data make similar inferences. Second, wherever contradictory or surprising outcomes occur, analysts must question their frameworks and the interpretation of available evidence in an effort to improve the coherence between theory and data. Finally, analysts must demonstrate a commitment to explore and evaluate rival theories in order to ensure that they do not isolate themselves and lose contact with alternative perspectives. Most importantly, Thagard's emphasis on scientific behavior being contingent on the social circumstances and connectivity of different groups suggests that the standards for scientific intelligence analysis should be contingent: Scientific status is gained and lost based on how the intelligence community develops its expertise and capabilities to analyze problems over time, and is not determined by a single analytic product or estimate.

This examination of demarcation criterion and alternative characteristics of science demonstrates how many of the traditional comparisons between science and intelligence used by the intelligence community have vastly oversimplified their differences and concealed more than they reveal about intelligence as an intellectual and organizational activity. While few intelligence problems fit the rigid logical structure advocated by Popper, neither does much of science. Ziman, Kuhn, Lakatos, and Thagard each expanded the definition of science, and shifted the demarcation criteria from simple

logical argumentation to more complex social processes and community structures. In doing so, they provided many different lenses for analyzing the intelligence community and have introduced criteria that include considerations of the psychological, social, and historical circumstances of intelligence communities and activities. Thus, differentiating between intelligence as a science or an art mischaracterizes the problem facing scholars and practitioners alike, given the diversity of beliefs about science and scientific processes.

Science, Time, Induction, and Intelligence

Alternative perspectives on demarcation suggest that science has far more in common with intelligence analysis than many in the community believe. This fact alone, however, does not make science a suitable model for intelligence. Ben-Israel argued that relating science and intelligence required differentiating between science as a methodological process for developing knowledge, and employing the substance of scientific theories under the guise of scientism. As he noted, “We have to adopt the logic of the scientific research process while ignoring its specific content,” arguing that intelligence analysts may need to develop their own theories and methods of analysis that are distinct from their peers in the scholarly community, but nevertheless conform to scientific standards.³⁸³

In spite of their differences, all of the philosophical ideas discussed earlier viewed the scientific research process as evolutionary—the implications of which are

³⁸³ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 689.

significant.³⁸⁴ Evolutionary processes are historical—each generation builds upon the successes of past generations, and adapts as a result of the challenges that were previously encountered in response to selection pressures—thus evolution is essentially a narrative of survival and the transition of features across time, whether those features are genes, beliefs, or problem-solving heuristics. Popper, Ziman, Kuhn, Lakatos, and Thagard all provided different versions of how theories survived within social communities when confronted with validating or falsifying evidence. Additionally, each of these philosophers agreed that science was ultimately an indicative process where the role of theory was to explain phenomena found in the real world, and that empirical observation was the ultimate arbiter of a theory’s viability. This provides a critical point of departure between science and intelligence based on their respective relationships between time and knowledge.

A Priori, a Posteriori, and the Temporal Foundation of Science

Most scientists and members of the intelligence community divide inferences between deductive and inductive claims. Implicit in this division is the role of time as an organizing concept around which distinctions between the imaginative world of logic and the empirical world of observation are made. A more fruitful approach from the perspective of intelligence is to organize inferences or conjectures around when they are made, specifically between what is known before and after collection occurs and experience is gained. Thus, what can be known beforehand, or *a priori*, belongs to the

³⁸⁴ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 676-678.

realm of deduction while what is learned after the fact, or *a posteriori*, belongs to the realm of induction and owes its claims to experience—both observation and experimentation.³⁸⁵

Shifting the terminology away from deduction and indication to *a priori* and *a posteriori* knowledge reveals a central problem in the treatment of intelligence as science. Specifically, how can intelligence professionals peer into the future, past the limitations of their observations, and still behave scientifically? Such a framing of the problem returns to the epistemological framework advanced by Kent, which divided intelligence into those matters known from experience and observation—basic and current intelligence—and those that are conjectures about the unknown and unseen—estimative or speculative intelligence.³⁸⁶

In science, a reasonable response to any conjecture about the future is simply to allow nature to run its course and wait for suitable opportunities to observe desired phenomenon, or perform an experiment in order to generate observational data. Yet, this is rarely an option for intelligence analysts and the policy makers they support. Intelligence analysts might rely on deduction by building assessments on top of their assumptions according to the rules of logic. Estimates of this kind may be internally consistent, but lack any sense of whether or not they are sound or relevant without any contact with the real world.

³⁸⁵ Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 25-30, and 234.

³⁸⁶ Sherman Kent, "Estimates and Influence," *Studies in Intelligence*, Vol. 12, No. 3 (1968), pp. 11-21.

A more fruitful and commonly applied analytic approach grounds estimates in available information and experience, and thus resembles inductive science with a major caveat: The cycle that closes the loop between theory and observation remains open and incomplete. As a result, limiting intelligence analysis to what can be known *a priori* constrains the ability for analysts to identify and correct inferential errors whose presence is not, and sometimes cannot, be known in advance. Thus, intelligence and science may employ similar methods, but their underlying epistemology is different due to the truncation of the scientific method. Indeed, Richard Danzig noted that policy making under the conditions of uncertainty was akin to “driving in the dark,”—in which case inductive science may be considered “driving while looking in the rearview mirror.”³⁸⁷

The *a priori* constraint that confines intelligence to deduction or unchecked inductive inference denotes a stark contrast between analysis and science, but overstates the real-world limitations of intelligence communities. Although the inductive loop that characterizes the scientific method may be incomplete when employed by the intelligence community, it is not non-existent. Intelligence estimates are complex networks of interrelated hypotheses, theoretical concepts, measurement schemes, etc.—they are rarely verified nor falsified completely by single observations. While analysts may be unable to wait and compare their judgments with reality, they may be able to test and evaluate parts of their analytic judgments over time, assessing the relative strengths and weaknesses of particular assumptions and inferences with each new observation. However, whereas

³⁸⁷ Richard Danzig, *Driving in the Dark: Ten Propositions About Prediction and National Security* (Washington, DC: Center for New American Security, October 2011); and Mark Palmer, *Gain Capital: A Glimpse in the Future of Analytics?*, May 4, 2012, http://streambase.typepad.com/streambase_stream_process/2012/05/inside-gain-capitals-use-of-streambase-liveview.html (accessed on June 23, 2012).

scientists may be capable of decomposing their theories into individually testable hypotheses, intelligence analysts assessing complex, open systems may not be able to do so. As a result, ambiguity will likely persist regarding the relationship between evidence and inference regarding precisely what the implications of new information means to the larger analytic assessment.

Given the incomplete character of estimates that must be generated *a priori*, it is useful to consider some of the conditions under which inductive inference fails, particularly when denied systematic opportunities to gather information in order to falsify their claims. Intelligence communities must deal with three significant challenges—dynamics, asymmetric inferences, and counterfactuals—each of which violate the foundations of inductive inference and empiricism, and yet cannot be avoided as they reside at the intersection between intelligence and policy. Before examining these three issues, however, two additional features of science must be examined in order to clarify how and why gaps between science and intelligence can be closed in spite of their differences.

Laws and Theories: Description vs. Explanation

Discussions of *a priori* vs. *a posteriori* knowledge beg questions about the nature of prediction, and its justification. Science may focus on description, identifying and classifying patterns and relationships, or science may focus on explanation, where mechanisms and processes are theorized in order to suggest how and why patterns and relationships exist. Each approach may generate predictions, but these predictions are of distinctly different kinds.

Understanding the basis of predictions requires differentiating between laws and theories. Laws are often referred to as facts about the world that are stable and confidently known, while theories are viewed as tentative and unreliable conjectures. For example, creationists' challenges to evolution have argued that evolution is just a theory, but not a fact. Such statements, however, are misleading. Laws are facts in the sense that they are descriptions of empirical phenomenon—they answer questions of *what*, and ultimately rest on the occurrence of correlations in empirical data.³⁸⁸ For example, the law of gravity refers to the fact that objects exert force on one another in proportion to their relative mass and distance. Yet, the law of gravity does not explain *why* objects attract one another: Answering such a question is the role of theory.³⁸⁹

[Theories] are neither true nor false. Theoretical notions find their justification in the success of the theories that employ them. Of purported laws, we ask: "Are they true?" Of theories, we ask: "How great is their explanatory power?" Newton's theory of universal gravitation provided a unified explanation of celestial and terrestrial phenomena. Its power lay in the number of previously disparate empirical generalizations and laws that could be subsumed in one explanatory system, and in the number and range of new hypotheses generated or suggested by the theory, hypotheses that in turn led to new experimental laws....

Laws are "facts of observation"; theories are "speculative processes introduced to explain them." Experimental results are permanent; theories, however well supported, may not last... Laws remain, theories come and go. Theories, then, explain what laws describe. Scientists' efforts to derive causal, explanatory theories from facts have proven both difficult in practice, and establishing the proper boundaries between description and explanation has remained a contentious issue within philosophy.³⁹⁰

³⁸⁸ Erwin Schrodinger, *What is Life?* (New York, NY: Cambridge University Press, 2004), pp. 3-18.

³⁸⁹ Dennis Overbye, "A Scientist Takes On Gravity," *The New York Times*, July 12, 2010, <http://www.nytimes.com/2010/07/13/science/13gravity.html?pagewanted=all> (accessed on March 24, 2012).

³⁹⁰ Kenneth N. Waltz, *Theory of International Politics* (Boston, MA: McGraw Hill, 1979), p. 6.

By distinguishing between laws and theories it is easier to understand the ways in which inductive inferences may be developed in rather peculiar ways, particularly as they relate to the use of statistics and generalization. Doing so allows for predictions to be generated from descriptions without providing explanations. Thus, statistical frequencies may become predictive probabilities, and yet never explain why past patterns and relations exist or should be expected to persist in the future.

Alternatively, theories could be employed to generate predictions, but doing so has proven to be difficult in practice. It is often assumed that if explanations for how a system has behaved in the past are available, they can be used to predict the future of a system. This, however, has proven problematic in practice given that explanations are often asymmetric.

The asymmetry of explanation is seen in the classic flagpole problem in philosophy, in which a flagpole of a known height casts a shadow of an unknown length. By knowing the height of the flagpole and the position of the sun relative to the Earth, the length of its shadow can be predicted and explained based on orbital mechanics and the physical properties of light. Light and position explain *why* the shadow is a given length. The same body of knowledge about the Earth's orbit and properties of light can be inverted to predict the height of the flagpole if only the length of its shadow is known. However, the second prediction provides no explanation as to *why* the flagpole is a given height—it is therefore a non-explanatory prediction. The effect of casting the shadow

can be employed to predict its cause, the flagpole, but cannot explain why the flagpole is the height that it is or its location.³⁹¹

In many sciences, theories can provide richly detailed explanations of change over time but cannot project forward beyond making the most general of predictions that are often overly vague when applied to policy. The theory of evolution may explain the length and shape of finches' beaks, or the correlated changes in the shapes and smells of wasps and orchids, but cannot predict how these features of populations will change over time.³⁹² In contrast, predictions based on descriptive statistics are often highly accurate but dependent upon the assumptions of homogeneity and stationarity discussed below. As a result, inferences based on statistical methods provide an alluring but ultimately unsatisfying solution to many strategic intelligence problems.

The intelligence community is often chastised for failure to predict events, yet the credibility of their assessments is determined by whether consumers believe they provide explanations as to why and how events might unfold. As Ben-Israel noted;

The intelligence researcher seeks to answer *why-questions* (or *how-questions*), not *what-questions*.... Questions such as, "*What happens near the Canal?*" or "*What are the Syrians doing?*" are, of course, important, but their importance is linked to the fact that the answers can constitute a data-base which will answer the *more* important "why-questions"....

To answer "why-questions" ("why is P?") we have to supply *causal explanations*: the general formula of a why-question answer is

³⁹¹ Samir Okasha, *Philosophy of Science: A Very Short Introduction* (New York, NY: Oxford University Press, 2002), pp. 41-47; Peter Godfrey-Smith *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 191-195; and Curtis Brown, *Philosophy of Science: Explanation*, January 24, 2011, <http://www.trinity.edu/cbrown/science/explanation.html> (accessed on March 24, 2012).

³⁹² Peter R. Grant and B. Rosemary Grant, *How and Why Species Multiply: The Radiation of Darwin's Finches* (Princeton, NJ: Princeton University Press, 2008); and John N. Thompson, "Coevolution," in Simon A. Levin, ed., *The Princeton Guide to Ecology* (Princeton, NJ: Princeton University Press, 2009), pp. 247-252.

“because of Q.” We should, therefore, be able to show that Q causes P (formally, that $Q \rightarrow P$), thus having a *theoretical* understanding of the processes in question. To gain this understanding, how-questions (“How do the Syrian military forces deploy to attack?”) must be answered.³⁹³

Statistical approaches may appear to provide analysts with opportunities to generate accurate estimates, but in actuality fail to satisfy consumers who are searching for causes around which to mitigate threats and capitalize on opportunities. For example, the rise of the “big data” movement in science, and search for collective intelligence based on the “wisdom of crowds” have made inroads into the intelligence community with the hope of improving the predictive accuracy of analytic estimates and products.³⁹⁴ In each case, the community’s impulse for dealing with uncertainty is to gather more data, and in doing so, become increasingly dependent on the development of non-explanatory predictions—whether resulting from data mining, machine learning, prediction-markets, etc.—and inadvertently induce failure by overwhelming analysts with irrelevant and potentially misleading information.³⁹⁵

Reconsidering the relationships between description and explanation, and laws and theories, suggests that these efforts will not serve the intelligence community as well as advocates wish. The rendering of estimates based on statistical regularities found in data may simultaneously yield increasingly accurate predictions that cannot serve the

³⁹³ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 701.

³⁹⁴ James Surowiecki, *The Wisdom of Crowds* (New York, NY: Doubleday, 2004); Michael Nielsen, *Reinventing Discovery: The New Era of Networked Science* (Princeton, NJ: Princeton University Press, 2011); and David Weinberger, *Too Big to Know: Rethinking Knowledge Now That the Facts Aren’t the Facts, Experts are Everywhere, and the Smartest Person in the Room in the Room is the Room* (New York, NY: Basic Books, 2012).

³⁹⁵ Michael H. Handel, “Strategic Surprise: The Politics of Intelligence and the Management of Uncertainty,” in Alfred C. Mauer, Marion D. Tunstall, and James M. Keagle, eds., *Intelligence: Policy and Process* (Boulder, CO: Westview Press, 1985), p. 246.

needs of consumers. For example, when examining the prospective uses of prediction markets for intelligence analysis, Pherson and Heuer noted:

Prediction Markets do have a string record of near-term forecasts, but intelligence analysts and their customers are likely to be uncomfortable with their predictions. No matter what the statistical record of accuracy with this technique might be, consumers of intelligence are unlikely to accept any forecast without understanding the rationale for the forecast and the qualifications of those who voted on it.³⁹⁶

As long as intelligence questions and policy makers' concerns remain focused on questions of *how* and *why*, many predictive methods and the analytic tradecraft are likely to be misaligned. Unless descriptive methods for generating predictions are paired with theoretical analyses, they are unlikely to make significant contributions to the policy process and leave intelligence producers isolated from consumers.

Distinguishing between laws and theories clarifies the bases upon which intelligence assessments rest. Predictions based on projections of statistical patterns, or the operations of scientific laws, are fundamentally different from those that rest on the causal mechanisms and processes. Failure to differentiate between the two can often lead to intelligence applications that mischaracterize the problems analysts face, squandering collection resources, investments in technology, and human capital. Absent any understanding of causal explanation regarding how and why foreign actors and international systems behave as they do, non-explanatory predictions may leave policy makers impotent and unable to assess the consequences of the choices that they, or others, might make.

³⁹⁶ Richards J. Heuer, Jr. and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly, 2011), p. 210.

Generalization

Thus far, discussions of demarcation criteria, paradigms, research programs, induction, and laws have all pointed towards one of science's most important epistemic values: the search for general, universal knowledge. Theories are improved as the number of problems and puzzles they solve expands. Inductive patterns achieve the status of scientific law as they are observed across increasingly large and diverse numbers of cases. In each case, the narrower and more specific the theory or pattern, the less impressive the scientific result.

The problem posed by generalization is one that is deeply seated in scientific practices, and has often been employed as a type of demarcation criterion in its own right. For example, clashes between political scientists and historians are often based on the extent to which each is comfortable looking for or believing in general patterns at the expense of uncovering and examining idiosyncratic differences across cases. For example, Gaddis noted that while political scientists emphasized the development and testing of general theory, historians strongly resisted this approach in favor of making claims about particular cases.

...the best reason of historians' distrust of theory is that we sense in it a lurking Catch-22: theorists seek to build universally applicable generalizations about necessarily simple matters; but if these matters were any more complicated their theories wouldn't be universally applicable. From our perspective, then, when theories are right they generally confirm the obvious. When they move beyond the obvious they're usually wrong. Historians proceed quite differently: we construct narratives in order to reconstruct complex events. In doing so we subvert as much as we

facilitate generalization. Because life is complicated, so is history.³⁹⁷

Generalization often plays a subtle role in analysis, and can subtly shift the attention of analysts from particular, unknown cases back to prior experiences with which producers and consumers are more familiar. While generalization is often logical, even unavoidable, it can lead to major analytic problems and failures when unrecognized, as was the case in evaluating Iraq's WMD capabilities discussed earlier.³⁹⁸ By assuming the existence of a general pattern of behavior regarding nuclear disarmament, the analysis of Iraq's WMD was not based on the particular circumstances of Iraq, i.e. its strategic calculations and decision making, political leadership, domestic and regional conditions, or technical capabilities. Instead, it depended on generalizations about the choices of South Africa, Ukraine, and Kazakhstan. As a result, the conclusion of Iraqi deception only considered the specifics of Iraqi behavior to the extent that it did not match generalizations from other cases, rather than alternative motivations based on its unique strategic context. Additionally, assuming that a pattern based on three cases constituted a scientific law upon which a confident assessment could be made (hence the employment of a proof by contradiction), reveals the extent to which the characteristics of good science and the pursuit of generalization can be detrimental to the needs of analysis and its emphasis on specifics.

³⁹⁷ John Lewis Gaddis, "In Defense of Particular Generalization: Rewriting Cold War History, Rethinking International Relations Theory," in Colin Elman and Miriam Fendius Elman, eds., *Bridges and Boundaries: Historians, Political Scientists, and the Study of International Relations* (Cambridge, MA: MIT Press, 2001), p. 302. Also see John Lewis Gaddis, *The Landscape of History: How Historians Map the Past* (New York, NY: Oxford University Press, 2004).

³⁹⁸ Condoleezza Rice, "Why We Know Iraq is Lying," *The New York Times*, January 23, 2003, <http://www.nytimes.com/2003/01/23/opinion/why-we-know-iraq-is-lying.html?pagewanted=all&src=pm> (accessed on March 22, 2012).

Intelligence Challenges to Science

Differentiating between theories and laws explains why the normal inductive version of science is inappropriate for intelligence analysis. Likewise, many of the epistemic values held by scientists are problematic for intelligence and policy analysis more broadly, e.g. generalization and parsimony. This final section examines two challenges to intelligence as science—dynamics and counterfactuals—which render intelligence questions problematic for many conventional social science research methods.

Dynamics

The foundation of induction within science rests on the belief that past experience provides the basis for future expectations. Implicit in the structure of inductive inference are the assumptions that systems are stationary and homogeneous, both of which are difficult to sustain in the context of the international system and the challenges intelligence analysts face when coping with active, willful rivals.³⁹⁹

The assumption that a system is stationary implies that whatever processes and relationships existed in the past will continue into the future. Thus, if a correlation is found between two variables in the historical record, the association will persist.⁴⁰⁰ This allows for the identification of general laws. Stationary systems also allow for the conversion of frequencies into probabilities, again relying on the assumption that

³⁹⁹ Alan Beyerchen, "Clausewitz, Nonlinearity, and the Unpredictability of War," *International Security*, Vol. 17, No. 3 (Winter 1992-1993), pp. 66-67.

⁴⁰⁰ Thomas A. Bass, *The Predictors* (New York, NY: Holt, 1999), p. 204.

statistical descriptions of observed populations will accurately predict the features of similar populations in the future.

The assumption of homogeneity implies that cases within a sample are identical in all relevant dimensions, meaning that they can be handled as undifferentiated units. While such an assumption may be justified in the case of physical particles such as electrons, where each is expected to behave exactly as the others, it is problematic when thinking about human beings in social systems because each unit may act based on its unique experiences, beliefs, and goals.

The mindfulness of social agents differentiates them from physical agents. Social agents often have mental models that they use to inform their behavior. Moreover, unlike physical agents, there is a plasticity in social agents who can change how they behave if outcomes are not to their liking. The rates and mechanisms of change may well depend on the system. An individual human agent can conduct a “thought” experiment and rapidly alter her behavior, while a lower-level biological agent may be destined for much slower changes via less direct mechanisms like natural selection. In contrast, bosons, quarks, electrons, and atoms, at least as far as we know, cannot change their rules modulo quantum fluctuations.⁴⁰¹

If systems are homogeneous, then numerous targets of intelligence and concerns of consumers are for naught. For example, do the particular personalities of foreign leaders, organizational structures of governments, and unique historical and cultural experiences matter? Or should rival states simply be regarded as functionally undifferentiated units whose individual goals can be assumed to be identical?⁴⁰² As outcomes within the international system, whether historical or potential, are seen to rest on more idiosyncratic factors, the assumption of homogeneity becomes less useful.

⁴⁰¹ John H. Miller and Scott E. Page, *Complex Adaptive Systems: An Introduction to Computational Models of Social Life* (Princeton, NJ: Princeton University Press, 2007), p. 100.

⁴⁰² Ken Waltz, *Theory of International Politics* (Reading, MA: McGraw Hill, 1979), pp. 97, and 143-144.

Assumptions regarding stationarity and homogeneity encourage generalization, yet they are fragile in cases where systems and their actors are dynamic. If the relationships between actors change over time and the subjects of observations display heterogeneous properties and behaviors, then the entire basis of inductive science may be called into question on practical and epistemological grounds. Indeed, inductive inferences fail when the past and future become disconnected, and previously identified patterns and relationships upon which estimates rest, no matter how stable, cease to exist or characterize the phenomenon of interest.⁴⁰³ Therefore, once systems are viewed as dynamic and historical, the bases of inductive inference are weakened, challenging the underpinning of science as it applies to intelligence analysis.

Counterfactuals

The primary challenge facing intelligence analysts and the policy makers they support rests on anticipating and characterizing potential futures and the consequences of alternative choices. What is rarely considered is whether analytic products that are highly situated and focused on serving the needs of particular decision makers in specific contexts can be evaluated empirically. In simple terms, how can analysts and decision makers know whether alternatives to the choices they made would have turned out for the better, the worse, or differently at all? The treatment of intelligence problems as one of generating and evaluating counterfactuals reveals several important features that are rarely considered in the broader context of science and analytic tradecraft.

⁴⁰³ Peter Godfrey-Smith, "Induction, Samples, and Kinds," in Joseph Keim Campbell, Michael O'Rourke, and Matthew H. Slater, eds., *Carving Nature at Its Joints: Natural Kinds in Metaphysics and Science* (Cambridge, MA: MIT Press, 2011), pp. 33-52.

Counterfactual claims are fundamentally causal and rely on the identification of mechanisms that link processes and outcomes. To suggest that a different choice will produce a different outcome implies that some causal process is affected by individual and collective decision making and actions, e.g. investing in one weapon system vs. another causes a different outcome to occur, such as the possession of greater combat power, more credible coercive threats, and a change in the balance of power.⁴⁰⁴ Whereas science has primarily deferred to evidence over theory, i.e. used empirical observations as the adjudicator of theoretical disputes, counterfactuals assign primacy to theory over evidence in the analytic process.

Counterfactuals also admit a more complex and nuanced understanding of the empirical record once probabilistic elements are admitted. Given that the empirical record provides a single instance of the case in question, precisely how should observations be used to support or refute an analytic estimate? If one assessment says the likelihood of an outbreak of war is low and another says it is high, how should evaluators of intelligence determine the quality of analysis based on the occurrence or nonoccurrence of war, if neither assigns likelihoods of 0 or 100 percent? Which assessment, and implied theory that motivated it, is better supported by the evidence? If war occurs, is it possible to know if the assessment that argued the event was unlikely but possible is truly worse than the assessment that determined war was more probable?

⁴⁰⁴ Philip E. Tetlock and Aaron Belkin, "Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives," in Philip E. Tetlock and Aaron Belkin, eds., *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives* (Princeton, NJ: Princeton University Press, 1996), pp. 1-38.

If the empirical record is no longer viewed as providing a set of homogeneous independent cases that characterize the patterns and relations in a stationary system, history itself ceases to be an effective arbiter between competing theories as there is no empirical way to differentiate between worlds where the most likely events occurred from those where the least likely (but still possible) outcomes resulted. Simply put, if any historical event were replayed, would the identical result occur, and if not, how different might the results be?

...when you're working with multiple intersecting variables over long periods of time, the conditions that prevail at the beginning of a process guarantee very little about its end. "Alter any early event, ever so slightly," the paleontologist Stephen Jay Gould has written of his field, "and the evolution cascades into a radically different channel." This is not to say that the history of life—or, by implication, history in general—lacks patterns: "the divergent route...would be just as interpretable, just as explainable after the fact, as the actual road. But the diversity of possible itineraries does demonstrate that eventual results cannot be predicted at the outset."⁴⁰⁵

Finally, counterfactual considerations reveal the problem of agency and structure in complex social systems.⁴⁰⁶ As Philip Tetlock, Richard Ned Lebow, and Geoffrey Parker noted when discussing counterfactuals related the rise of West in world history, counterfactual assessments depend on and expose deeply held beliefs about the social world and the role of individuals within it.

Surveys of professional historians have shown that observers who lean toward the political Right are more likely to maintain that things had to work roughly as they did and that Western dominance has been in the historical cards for a long time (sometimes as far back as a thousand

⁴⁰⁵ John Lewis Gaddis, *The Landscape of History: How Historians Map the Past* (New York, NY: Oxford University Press, 2002), p. 66.

⁴⁰⁶ Philip E. Tetlock, Richard Ned Lebow, and Geoffrey Parker, "Unmaking the Middle Kingdom," in Philip E. Tetlock, Richard Ned Lebow, and Geoffrey Parker, eds., *Unmaking the West: "What-If?" Scenarios That Rewrite History* (Ann Arbor, MI: University of Michigan Press, 2006), pp. 1-13.

years). Insofar as these observers tolerate explicit counterfactuals at all, they favor second-order counterfactuals—which bring history back in fairly short order—that concede that, yes, this or that surface cause could have taken on a different value and rerouted events briefly, but deeper forces would have returned history to something much like the trajectory we are now on. To these scholars, the West achieved geopolitical dominance because it exemplified distinctive cultural values and possessed unique political assets that conferred a long-term competitive advantage in creating and applying new technologies. The West won because it got certain things right—displaying more respect for property rights, implementing a clear separation of church and state, granting greater freedom to launch independent inquiry—that the rest got wrong. For them, any attempt to imagine counterfactual scenarios in which Western primacy is easily undone by minor twists of fate—a botched assassination here or a delayed invention there—will fail for the simple reason that the roots of the success of the West and of the failure of the rest lie deeply embedded in the mores, folkways, and institutional habits of the relevant societies.

By contrast, observers on the political Left tend to deride such thinking as “triumphalist.” They find the rise of Europe and then of North America to global dominance during the last five hundred years as just as improbable.... For them, the rise of the West was an accident of history, and Western hegemony a fluke, a one in a million shot that can be readily undone—at least in our imaginations—by altering minor background conditions as late as the seventeenth or eighteenth centuries: if a key individual had died slightly earlier or later, or if the weather had been cloudier or windier, “we” would find ourselves in a very different world. These scholars also deny that there is anything superior about Western culture when it is compared with the spirituality and communal solidarity of many African and Asian societies.⁴⁰⁷

A similar point has been made by Philip Sabin, who has examined the uses of simulation and war games in the study of history, particularly for adjudicating between competing and often incompatible accounts of ancient warfare.⁴⁰⁸ Sabin noted that many of the most consequential choices in simulation design touch directly on alternative views

⁴⁰⁷ Philip E. Tetlock, Richard Ned Lebow, and Geoffrey Parker, “Preface,” in Philip E. Tetlock, Richard Ned Lebow, and Geoffrey Parker, eds., *Unmaking the West: “What-If?” Scenarios that Rewrite World History* (Ann Arbor, MI: University of Michigan Press, 2009), p. 8.

⁴⁰⁸ Philip Sabin, *Lost Battles: Reconstructing the Great Clashes of the Ancient World* (New York, NY: Continuum, 2009).

of counterfactuals—where simulation developers must determine just how much players’ choices can affect outcomes. This approach to game design and simulation reflects different theories of history with respect to the relationship between structure and agency. The stronger the role of structure, the less the decisions, actions, and interactions of the actors can alter outcomes, while the stronger the role of agency, the more contingent outcomes become. These alternatives are shown in Figure 4-2 below.

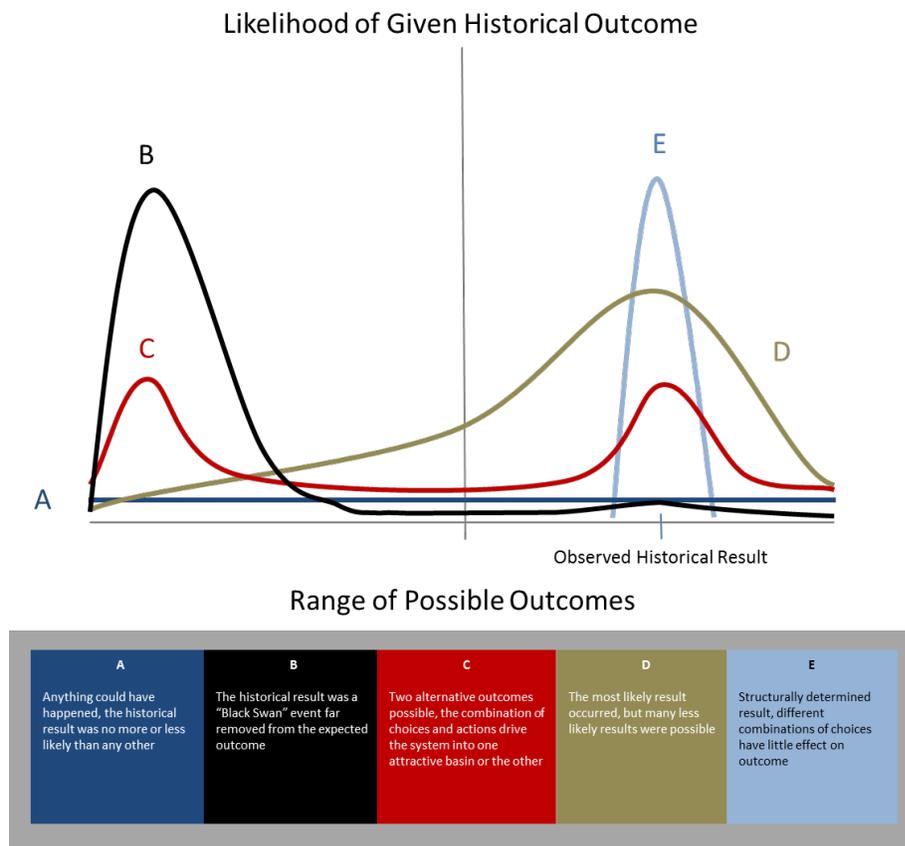


Figure 4-2: Agency and Structure in Counterfactuals. The figure above shows different beliefs about how historical outcomes may have been structurally determined or resulted from agent-level choices and actions. The x-axis depicts the range of different possible outcomes, while the y-axis shows the likelihood of an outcome occurring. Each colored line shows different philosophical beliefs about the nature of history and the reachability of different results. Figure 4-2 is adapted from Philip Sabin, *Simulating War: Studying Conflict Through Simulation Games* (New York, NY: Continuum, 2012), p. 56.

Science's commitment to generalization has biased the development of theories and research methods toward structural assessments and explanations, thus downplaying the consequences of individual's choices, personalities, and other features that are not easily compared across cases or preserved in the historical record. Because of the (assumed) stationary and homogeneous character of physical systems, such microlevel concerns were often set aside in physics and chemistry, allowing for the statistical mechanics of extremely large populations to alleviate the burden of examining individual units in detail. In the social sciences (and an increasing number of problems in physical and natural sciences), such assumptions, and the employment of the analytic methods that rely on them, are increasingly seen as problematic. For example, large-N studies enable cross-case comparisons on precursors to international war, state failure, or economic growth, yet they cannot reconstruct the particular perceptions (and misperceptions) and choices of the principals and agents whose decisions produced or prevented particular outcomes. As a result, it may be possible to reach conclusions on the effects of economic growth, international trade, and monetary policy on Germany's propensity to go to war following the interwar period by comparing 20th century Europe's structural conditions with those of other international systems at other times, while it may be impossible to empirically assess the role of individual leaders' personalities, beliefs, and choices, e.g. Hitler's personal contribution to Germany's rise and international relations in the interwar period, because there are no opportunities to observe interwar Germany with a different head of state.

The problem of counterfactual analysis pits the central premises of science as an inductive, empirical venture, producing general knowledge in conflict with the needs of intelligence analysts and policy makers attempting to choose between dimly lit, perhaps unknowable futures. Philip Bobbitt summarized this problem when discussing the formulation of strategy as follows:

Many commentators believe that the turning point in the 1980 U.S. presidential elections came in the first debate between the candidates when Governor Reagan asked the American people to consider the question, “Are you better off today that you were four years ago?” Indeed, this riposte was so successful that it was used in the 1984 debate by Reagan’s opponent, Walter Mondale; and used again by George Bush against Michael Dukakis; and then used by Governor Clinton against President Bush.

Such a question, however, can scarcely be the measure of a presidential administration because the one thing we know is that things will never stay the same for the length of a presidential term, regardless of who is in power. Governor Reagan ought to have asked the public in 1980, “Are we better off now than we have been if President Ford had held office these last four years?” This is the measure of the choice to be made, which might be phrased: “Will we be better off in four years, not ‘than we are now’ but ‘because of the choice we are asked to make now’?”⁴⁰⁹

Policymaking is fundamentally an agent-level activity, where the notion that choices have consequences is the *sin qua non* of governance. As long as inductive science serves as the model for analytic tradecraft, analysts may be poorly served by implicit structural biases in the development, verification, and falsification of theories that can be empirically tested and generally applied. Absent attention to, and methods specifically tailored for, exploring agency and contingency, policy makers may well be stuck with assessments that tell them they have no choices to make because their

⁴⁰⁹ Philip Bobbitt, *The Shield of Achilles: War, Peace, and the Course of History* (New York, NY: Anchor Books, 2002), p. 10.

decisions lack the power to affect outcomes—exacerbating the problem of optimism and pessimism in producer/consumer relations discussed earlier. To insist on arming intelligence analysts with the tools of empirical, inductive science exclusively is to leave them ill-equipped to address what policy makers need most, a way of understanding the consequences of their choices and those of others.

A Microscience for Intelligence Analysis

The preceding sections examined alternative perspectives on the problem of demarcation in the philosophy of science, the essential role of time in inductive inference, and limitations of empiricism and generalization from the perspective of the intelligence community. The conclusions that follow provide a framework for an alternative science that would suit the needs of the intelligence community. Importantly, many of the parts of this new approach to science are being developed by the scientific community as microlevel data becomes increasingly available, computational power increases, and ideas about evolution, contingency, and indeterminism can now be studied experimentally, even if they remain outside the bounds of empirical observation.

This version of science that addresses the needs of the intelligence community might be called a “microscience,” which explores individual, particular events and cases through the use of theory and data, but is not focused on the development of general knowledge or discovery of universal truths. Instead, microscience emphasizes linkages between microlevel, local circumstances, perspectives, contexts, and decisions, and macrolevel, systemic outcomes. Because the goals of microscience would be to uncover the features and dynamics of micro–macro linkages and agent–structure relations, many

of the popular approaches to scientific inquiry would be diminished as aids to addressing the most pressing questions. For example, the employment of large-N statistical studies, and use of representative agents whose interests, features, and behaviors are characterized by the mean of larger, diverse populations would be treated as too coarse to provide policy-relevant information. Yet, microscience is not a small-N approach akin to case studies that rely on process tracing alone because its motivation is to understand the extent to which outcomes are contingent—and therefore requires a means for generating and assessing counterfactuals that may be very near or far from actual historical events. In microscience, the empirical record is no longer the standard against which inferences are judged, but is seen as a single path to a historical outcome that cannot be assumed to be the most likely or consequential result.

Although the notion of microscience seeks to identify a mode of intelligence analysis that can satisfy the desire for analytic tradecraft to be practiced scientifically, it challenges the prescriptions of scientifically minded members of the intelligence community. For example, Bruce's emphasis on epistemology and self-correcting practices argued strongly for the primacy of empiricism in analysis, but offered little guidance for how analysts should consider contingencies and future possibilities about which no data is available and observation may not be possible.⁴¹⁰ Likewise, Ben-Israel's view of science and intelligence analysis emphasized the use of universal laws as the basis of analytic conjectures, and warned against employing trends in the same way

⁴¹⁰ James B. Bruce, "Making Analysis More Reliable: Why Epistemology Matters to Intelligence," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp.171-190.

because they are local, historically contingent phenomenon, i.e. relationships that may be short-lived, spurious, and context dependent.

A law (be it a law of nature, social law or any other) is always described by a *universal statement*. Usually it has the form of “whenever certain conditions are fulfilled, certain phenomena are observed,” or “every time event A occurs, it is followed by event B.” Sometimes, existential statements *appear* to be universal. This is usually so with statements about “trends” or “processes.” For example, “The trend of American voters is toward conservatism” is a singular, *not* a universal statement. It asserts the *existence* of a singular phenomenon, confined to a certain society at a certain period. Hence, *a trend is not a law*; it is simply a description of a singular phenomenon. The American voter need not obey the trend cited above, as he must obey gravity. Trends describe “local” behavior, in space or time, of a certain system and, hence, unlike laws, trends can change from moment to moment.⁴¹¹

A microscience for intelligence suggests the need to do the opposite, however.

Because laws are correlations, distinguishing them from trends is to acknowledge a quantitative, not qualitative, difference. The difference between a law and a trend may simply be the frequency with which they reappear in the historical record or experimental data. Projecting from trends may indeed produce highly misleading predictions with respect to general phenomena and claims of universal application. However, if analysts seek to produce insights into the contingent features of specific cases, then conjectures based on extending local trends may be precisely what are called for in order to examine how particular cases may unfold over time.

Gaddis provided an alternative way to think about differences between laws and trends: by discussing historical continuities and contingencies. Together, this framing

⁴¹¹ Isaac Ben-Israel, “Philosophy and Methodology of Intelligence: The Logic of Estimate Process,” *Intelligence and National Security*, Vol. 4, No. 4 (1989), p. 703.

may lead to a more natural or intuitive way to think about intelligence problems and the ranges of futures that are attainable from any given description of a system.

By continuities, I mean patterns that extend across time. These are not laws, like gravity or entropy; they are not even theories, like relativity or natural selection. They are simply phenomena that recur with sufficient regularity to make themselves apparent to us. Without such patterns, we'd have no basis for generalizing about human experience: we'd not know, for example, that birth rates tend to decline as economic development advances, or that empires tend to expand beyond their means, or that democracies tend not to go to war with one another. But because these patterns show up so frequently in the past, we can reasonably expect them to continue to do so in the future. Trends that have held up over several hundred years are not apt to reverse themselves within the next several weeks.

By contingencies, I mean phenomena that do not form patterns. These may include the actions individuals take for reasons known only to themselves: a Hitler on a grandiose scale, for example, or a Lee Harvey Oswald on a very particular one.... What all of these phenomena have in common is that they don't fall within the realm of repeated and therefore familiar experience: we generally learn about them only after they've happened.

We might define the future, then, as the zone within which contingencies and continuities coexist independently of one another; the past as the place where their relationship is inextricably fixed; and the present as the singularity that brings the two together, so that continuities intersect contingencies, contingencies encounter continuities, and through this process history is made. And even though time itself isn't structured this way, for anyone who's stuck within time—and who isn't?—this distinction between past, present, and future is close to universal.⁴¹²

An analogy for microscience may be understood as the study of an oil drop expanding on a surface. By starting from a rich description of an individual case, seen as the initial drop on a landscape, and then tracing how small changes to its specification—whether agents, networks, interactions, or macroscopic conditions—alter outcomes, the character of the landscape is revealed as the oil spreads to nearby points. The primary

⁴¹² John Lewis Gaddis, *The Landscape of History: How Historians Map the Past* (New York, NY: Oxford University Press, 2002), pp. 30-31.

way of learning the landscape is based on the development of counterfactuals and alternative theories, both explanatory and descriptive, which generate new cases by modifying the features of previously known ones.

Microscience is also consistent with existing tradecraft because it rewards analysts who remain open to exploring alternatives. The more alternative data sources and interpretations are employed, and the larger the set of theories that are considered as the basis for projections, the richer the generated landscape. This approach is then similar to existing tradecraft, e.g. the emphasis on exploring alternative hypotheses and interpretations of data encapsulated in ACH, but employs synthetic data from models and simulations (formal and informal) in order to fill in the landscape at scales and levels that cannot be done relying on empirical cases alone.⁴¹³

The use of modeling and simulation enables intelligence analysis within the microscience framework. Traditional efforts to employ models in analysis have proven problematic, and while not without successes, their integration into analytic tradecraft has been difficult to achieve.⁴¹⁴ There are many good reasons that explain why formal modeling has experienced difficulties becoming a significant part of analytic tradecraft.

⁴¹³ Steven C. Bankes and James J. Gullugly, *Exploratory Modeling: Search Through Spaces of Computational Experiments* (Santa Monica, CA: RAND, 1994); and Robert J. Lempert, Steven W. Popper, and Steven C. Bankes, *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis* (Santa Monica, CA: RAND, 2003).

⁴¹⁴ Michael K. O’Leary, William D. Coplin, Howard B. Shapiro, and Dale Dean, “The Quest for Relevance: Quantitative International Relations Research and Government Foreign Affairs Analysis,” *International Studies Quarterly*, Vol. 18, No. 2 (Jun., 1974), pp. 211-237; Richards J. Heuer, Jr., ed., *Quantitative Approaches to Political Intelligence: The CIA Experience* (Boulder, CO: Westview Press, 1978); James B. Bruce, “Making Analysis More Reliable: Why Epistemology Matters to Intelligence,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 183-184; and Richards J. Heuer, Jr., and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly, 2011), pp. 22-23.

However, transitioning from mathematical formalisms to computational ones allows for dramatic increases in the complexity of what can be represented, addressing many of the problems produced by the assumptions of stationarity, homogeneity, rationality, and more, which have limited the contributions of formal models and simulations in the past.

Additionally, an approach to intelligence analysis that increasingly employs surrogate and artificial systems in order to evolve an understanding of the real world would keep analytic tradecraft consistent with the changing characteristics of science, which has become increasingly model-centric over time.⁴¹⁵

... a recognizable style of model-based science (self-conscious or not) may be increasing in prominence within science, on a decade-long and perhaps centuries long time-scale. This has eventually led to the development of a self-conscious practice of a model-driven form of science, guided by a distinctive set of ideals. This last development involves, among other things, a rejection of the idea that modeling is a mere heuristic adjunct [sic] to the real business of theory-construction. Instead, we have an embracing of the idea that models themselves can be the tools by which we represent and understand the world.⁴¹⁶

Thus many natural and social sciences are increasingly becoming experimental sciences, able to generate and explore alternative worlds through the use of simulations—a development that actually suggested an increasing alignment between scientific inquiry

⁴¹⁵ On the changing character of science and the role of models see Mary S. Morgan and Margaret Morrison, eds., *Models as Mediators: Perspectives on Natural and Social Science* (New York, NY: Cambridge University Press, 1999); Peter Godfrey-Smith, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003), pp. 230-231; Peter Godfrey-Smith, “The Strategy of Model-Based Science,” *Biology and Philosophy*, Vol. 21, (2006), pp. 725-740; and Daniela M. Bailer-Jones, *Scientific Models in Philosophy of Science* (Pittsburgh, PA: University of Pittsburgh Press, 2009).

⁴¹⁶ Peter Godfrey-Smith, “The Strategy of Model-Based Science,” *Biology and Philosophy*, Vol. 21, (2006), p. 730.

and analytic tradecraft.⁴¹⁷

Finally, microscience alters the ways in which systems are described. The epistemology of physics underwent a significant transformation with the development of quantum theory, which admitted and later required descriptions of systems that were probabilistic in nature. From the perspective of microscience, assessments of social phenomena and intelligence assessments may become increasingly incomplete if they are purely descriptive and only provide empirical information. Instead, more complete assessments and characterizations of systems and intelligence targets will provide a state description, as well as account for those aspects that are likely to remain continuous and those that are contingent on particular individual and collective choices.⁴¹⁸ Such products would be generated by a combination of rich, detailed, case-specific investigations, akin to the work of historians, journalists, and area specialists, and the development of models that simultaneously employ statistical, data-driven projections and dynamic, theory-driven simulations to explore and characterize potential futures.

⁴¹⁷ Joshua M. Epstein and Robert Axtell, *Growing Artificial Societies: Social Science From the Bottom Up* (Washington, DC: Brookings, 1996); Robert Axelrod, *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration* (Princeton, NJ: Princeton University Press, 1997); Joshua M. Epstein, *Generative Social Science: Studies in Agent-Based Computational Modeling* (Princeton, NJ: Princeton University Press, 2007); and John H. Miller and Scott E. Page, *Complex Adaptive Systems: An Introduction to Computational Models of Social Life* (Princeton, NJ: Princeton University Press, 2007).

⁴¹⁸ Robert Stalnaker, *Mere Possibilities: Metaphysical Foundations of Modal Semantics* (Princeton, NJ: Princeton University Press, 2011).

Chapter 5: The Schelling Segregation Model

This chapter examines several intelligence challenges using a very simple ABM of segregation dynamics originally developed by Thomas Schelling. The simplicity of the segregation model will allow for the arguments to focus on issues of model use, especially with respect to how ABMs can be employed address questions of individual agency in complex systems and the ways in which modeling and simulation can coordinate the efforts of analysts and collectors. When viewed from the perspective of analytic tradecraft, ABMs are shown to provide new research tools to assess questions of microscience discussed in Chapter 2.

Thomas Schelling's Segregation Model

During the late 1960s Schelling became interested in the problem of racial segregation in the US, and wondered whether the existence of racially segregated communities were indicative of the racial preferences of their inhabitants.⁴¹⁹ While this question may appear to be far removed from the concerns of contemporary intelligence analysts, it is quite relevant in two distinct ways. First, from a substantive perspective, many of the security and policy challenges that analysts must consider involve sectarian and ethnic conflict, and its potential to undermine the development of multiethnic, inclusive societies. Second, as a general question, Schelling characterized the problem of

⁴¹⁹ Thomas C. Schelling, "Dynamic Models of Segregation," *The Journal of Mathematical Sociology*, Vol. 1, No. 2 (1971), pp. 143-186.

racial segregation as a dynamic process that was not discernible from knowledge of individual preferences alone. As a result, he undermined a dominant paradigm or logic of intelligence analysis, suggesting how the way analysts and collectors approached intelligence problems could produce irrelevant or misleading judgments.

The Sectarian Segregation of Baghdad

The US invasion of Iraq and removal of Saddam Hussein’s Bathist regime significantly altered the social structure of Iraq’s sectarian communities. One of the most visible signs of the difficulties in building a new democracy could be seen in the rapid dissolution of Baghdad’s mixed communities and the rise of Sunni and Shia enclaves shows in Figure 5-1 below.

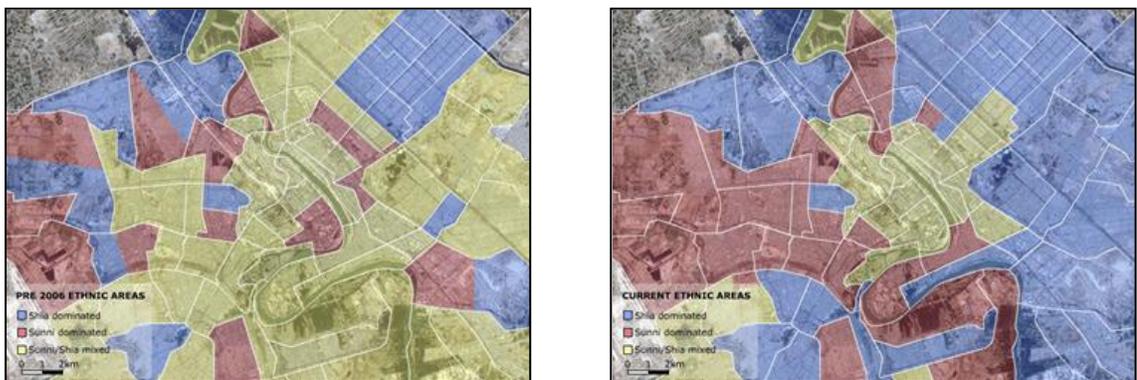


Figure 5-1: Baghdad’s Transition from Mixed to Segregated Communities. The map on the left depicts the sectarian mix of Baghdad prior to 2006, while the map on the right shows its mix after 2007. Shia neighborhoods are shown in blue, Sunni neighborhoods are shown in red, and mixed neighborhoods are shown in yellow. Sectarian maps of Baghdad taken from BBC, *Baghdad: Mapping the Violence*, http://news.bbc.co.uk/2/shared/spl/hi/in_depth/baghdad_navigator/ (accessed on May 8, 2012).

From the perspective of US policymakers, understanding whether changes in Baghdad's community structure and segregation are reflective of the population's preferences may be essential to determining whether Iraq has the potential to develop into a constitutional democracy or will descend into prolonged sectarian violence and civil war. If Baghdad's Sunni and Shia communities segregated as a result of the preferences of their members, then the prospects for Iraq's political and economic development may be hindered by persistent identity-based sectarian competition and conflict. Moreover, if its Sunni and Shia communities are viewed as the front in a larger regional or cultural conflict extending beyond national identities, then Iraq's neighbors may intervene covertly or overtly in its internal affairs, potential risking escalation into a destabilizing regional conflict.⁴²⁰ Alternatively, if Baghdad's neighborhoods have segregated as a result of social processes that are not reflective of individual preferences, then the loss of its mixed communities may be an ambiguous indicator of its political future. Thus, the issue of social segregation and the process by which it occurs is not one that intelligence analysts can assume lays beyond their responsibilities. Instead, it may prove crucial to providing context of the IC's most senior customers who must determine whether their policies are achieving their desired goals, whether new strategies are needed, or whether the objectives to which they have committed the nation's resources cannot be achieved.

⁴²⁰ The prospects of a transnational realignment of power within the Islamic world centered around the competition between Sunni and Shia identities and communities was most forcefully considered by Vali Nasr in his 2006 book, *The Shia Revival: How Politics within Islam Will Shape the Future* (New York, NY: W. W. Norton and Company, 2006). While Nasr argued for the increasing importance of sectarian identities in the Islamic world, others have examined the limits of Sectarian identities to fundamentally restructure the political and social order of the region given the continued importance of ethnic and national identities, as well as the historical fluidity of sectarian identities and the role of conversions and migration; for examples see Graham E. Fuller and Rend Rahim Francke, *The Arab Shi'a* (New York, NY: St. Martin's Press, 1999); Dale F. Eickelman and James Piscatori, *Muslim Politics* (Princeton, NJ: Princeton University Press, 2004); and Yitzhak Nakash, *Reaching for Power* (Princeton, NJ: Princeton University Press, 2006).

Complexity and its Challenge to the Intelligence Analysis Paradigm

Schelling's question regarding the relationship between individual preferences and the existence of segregated communities has proven to be a significant challenge to the dominant paradigm of intelligence analysis. Intelligence analysts and collectors have traditionally focused on uncovering the capabilities of foreign governments and assessing their intentions. Collection has traditionally targeted individuals with access to rival decision-making processes, or technical or bureaucratic artifacts, such as archives, communications, and weapons systems. Likewise, analysts have always sought to understand targets' behavior and decision making as goal seeking and problem solving, even if complicated by personal and bureaucratic politics, ideology, and other factors. Even the shifting of attention from state to non-state actors has not forced significant changes to the basic intelligence analysis paradigm. While the diversity of actors that intelligence professionals must understand has increased, dramatically expanding the goals, capabilities, and organizational profiles of targets, the fundamental strategic framework has remained consistent: Analysts and collectors remain focused on examining and understanding the capabilities and intentions of rivals, and operate under the assumption that these features provide the basis upon which assessments rest.

Schelling's concern for micro–macro linkages constituted a significant challenge to intelligence analysts and collectors because it reduced the consequences of individual strategic calculations and action.⁴²¹ The mere possibility that outcomes might not reflect

⁴²¹ This statement is not entirely precise. Individuals' actions may be calculated, rational, and strategic, but if their choices cannot produce the outcome they desire due to the ways in which agents interact, then knowledge of individual strategies may not provide a suitable basis for understanding how the system itself will behave.

the desires of individuals challenges the paradigm established by collecting and analyzing capabilities and intentions in favor of focusing on the processes by which individual actions aggregate. In cases where outcomes are emergent and result from the interactions between actors' choices and actions, increasingly detailed knowledge of individual capabilities and intentions may fail to produce meaningful intelligence assessments. Thus, Schelling's conjecture on the relationship between micro–macro linkages suggested how the structure of systems with complex, interdependent interactions undermined the intelligence community's dominant analytic paradigm. Further, Schelling's segregation model suggested the need to consider the fundamental epistemological limitations of analysis that could not be overcome by the most precise, in-depth and unambiguous descriptions of actors' capabilities and intentions. The mere possibility that systemic outcomes, such as segregated communities, may not result from the individual preferences of microlevel actors revealed the need for new analytic tools and theories within the social science and intelligence community.

The First Agent-Based Model

By framing questions about the emergence of segregated communities as the outcome of a dynamic process that aggregated the preferences and choices of individual agents, Schelling focused his research on the mechanism by which microlevel actions interacted. To explore this relationship, he created what is often regarded as the first ABM, originally examined by manually moving coins on a checkerboard since computing technology was not widely available at the time.

Schelling's model can be understood as a population of agents belonging to different identity groups, e.g. race, ethnicity, religion, class, etc., each trying to locate themselves amongst one another in space according to their individual preferences. In the simplest version of the model, each agent determines its happiness based on the composition of its eight neighbors as if it resided on the center of a tic-tac-toe board or lattice. This neighborhood structure is often referred to as a Moore neighborhood and is shown in Figure 5-2 below.

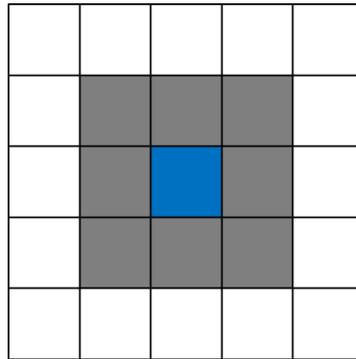


Figure 5-2: The Moore Neighborhood. The blue square above represents a single agent on a lattice. The surrounding grey boxes constitute the eight cells that depict its Moore neighborhood. In the Schelling Model, preferences of individual agents are expressed as the number of cells in the neighborhood that must be occupied by members from the same group in order for the agent to be happy and remain stationary.

The model's operation is based on a simple set of behaviors where agents observe their environment, determine their happiness, and decide whether they should move to a new location or remain in their current position. The underlying agent logic is as follows:

- Each agent counts the number of other agents in their Moore neighborhood that are members of the same group;
- If the count of similar agents is less than their preference, the agent will be unhappy. Likewise, it will be happy if the count of similar agents is equal to or higher than their preference;
- In a random order, all unhappy agents move to a random, empty location on the landscape;
- This process repeats until all agents are happy with their neighborhoods.

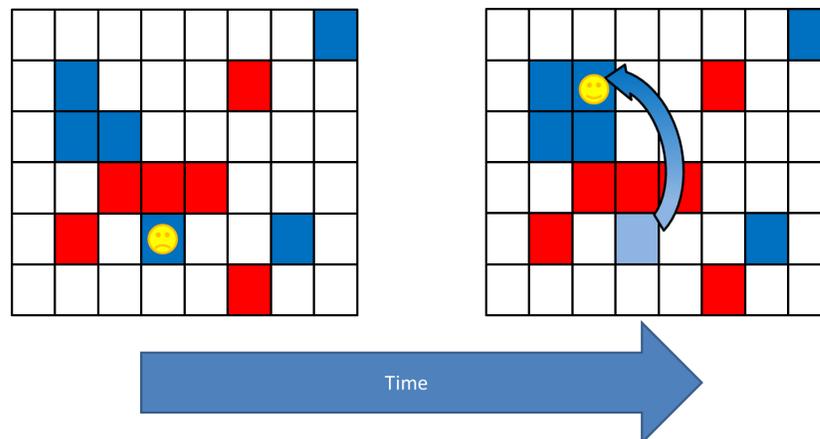


Figure 5-3: Segregation Model Step. The basis of the Segregation model is the movement of unhappy agents. In the example above, a single blue agent has no blue neighbors and determines that it is unhappy, and is marked with a yellow sad face. As a result, it moves to a random, empty location on the landscape, in this case into a group of three adjacent blue agents, creating a cluster or block of four agents arranged in a square. Assuming that it seeks to have a minimum of three neighbors with the same color, the agent would be happy in its new location. This process repeats until agents on the landscape are satisfied with their neighborhoods.

By animating populations of agents, where moves continue until all agents are happy based on their neighborhoods, communities with distinct spatial and ethnic or sectarian patterns are generated. As is often the case, the ways in which agents interact can produce systemic outcomes that are not discernible from their individual behavior and choices alone.

In the case of the segregation model, its results have proven surprising to many. Schelling found that representing agents' individual preferences as a fraction of the cells in their Moore neighborhoods with the same identity consistently generated segregated communities, even when the individuals were willing to be local minorities. More precisely, by expressing agent preferences as $n / 8$, where n was the desired number of neighbors of the same identity within in the agent's Moore neighborhood, the model consistently produced communities with high levels of segregation, particularly with respect to the value of n and the population density. Thus, for the case of $n = 3$, agents are highly tolerant and willing to be local minorities, remaining stationary and happy with as many as five of their neighbors being of a different identity. Yet the result is that a relatively small number of agents find themselves in mixed communities in spite of the overall level of tolerance within the population.

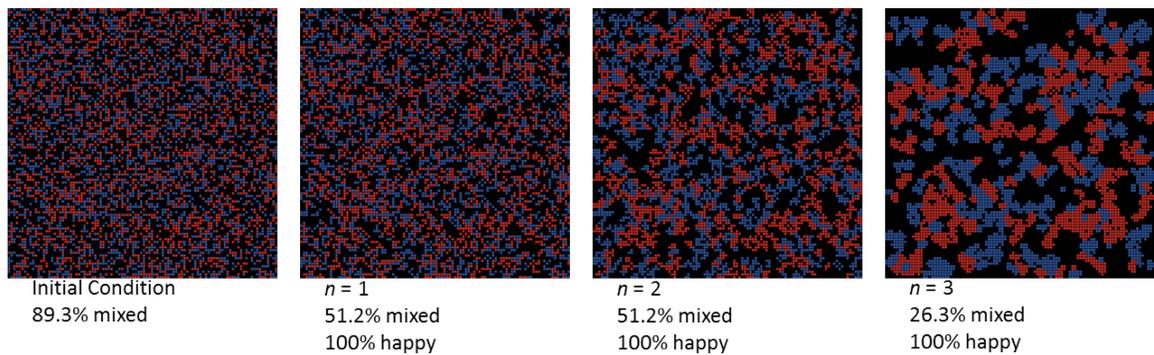


Figure 5-4: Segregation Model Initial and Final States. The above left panel shows the initial state of a population of five thousand agents representing two different identities shown as red or blue, with empty cells on the landscape in black. The three panels to its right show the different neighborhoods that can be generated for a given n , which represents the number of common neighbors each agent wants in order to be happy and stop moving. For each of the values of n shown [1, 2, 3] agents are willing to be minorities within their neighborhood, yet the percentage of agents residing in mixed neighborhoods declines rapidly, resulting in increasingly segregated communities.

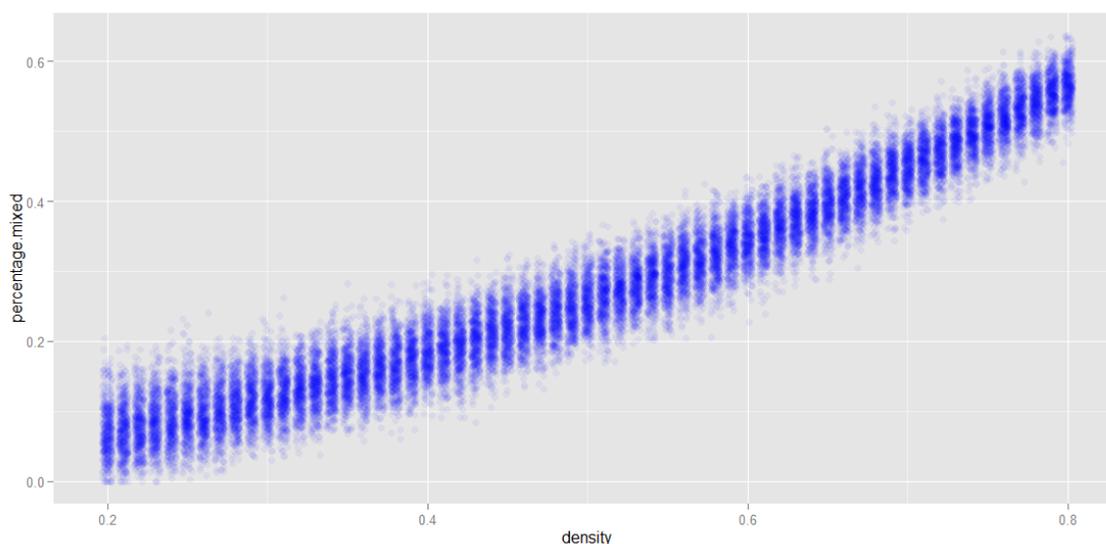


Figure 5-5: Segregation Model Response to Changing Population Density. The above plot shows the final percentage of agents living in mixed communities as a function of population density. Each point constitutes the result of a single run of the model for a given population density. All runs were performed on a 50 x 50 grid, and the identical conditions simulated 500 times for each population density over the interval of [0.20, 0.80] with a step size of 0.01. In all simulations, agent preferences were set to 3, and each group was the identical size. The graph shows that the percentage of agents residing in mixed communities increases with population density, as agents are less capable of finding large unoccupied regions on the landscape away from their respective out-group. Moreover, the variation in population density declined, as shown by the increasingly small vertical spreads of outcomes corresponding to higher densities.

Simulation results demonstrate how focusing exclusively on individuals and their preferences and actions may produce misleading or incomplete understanding of larger social circumstances. While the segregation model is quite simple, it nevertheless provides a useful pedagogical tool for intelligence analysts and methodologists seeking to understand how the dominant analytic paradigm may fail when addressing complex systems. The remainder of this chapter will continue to experiment with and develop the

segregation model in order to demonstrate how ABM can be integrated into intelligence analysis in order to address the challenges of microscience posed in Chapter 2.

Schelling Segregation Model as an Analytic Tool

The simple model that Schelling developed is simultaneously a toy model that offers a highly stylized and abstract description of multiethnic communities, geography, and social decision making. However, it also provides a powerful infrastructure upon which increasingly rich descriptions of social processes and decision making can be explored. Capitalizing on ABM as a research methodology in intelligence analysis requires that analysts and methodologists be patient and work through a development process that moves from simplistic abstractions to increasingly detailed and precise representations of complex social phenomena. While simplicity and parsimony are regarded as epistemic virtues in the scientific community, their value to analytic application and policy making is determined by more practical considerations. Starting computationally grounded intelligence analysis from simple, toy models is beneficial for the purposes of model verification and testing, and establishing a framework upon which more complex additions and revisions can rest. Thus, the virtues of Schelling's segregation models within the context of intelligence analysis are not its replication of real-world processes of community formation and change, but as a template upon which new features of interests to users can be incorporated.

Effectively employing ABM in intelligence analysis rests upon three interrelated factors:

- First, a sound theoretical understanding of the problem being investigated and the identification of mechanisms that relate the choices and actions of agents to one another and their environment;
- Second, a framework for experimental design that can generate the appropriate data needed for verifying the model's design and implementation, validate its configuration, and justify inferences of interest to analytic users;
- Third, a technical design of the model that simultaneously implements the model's theory accurately and allows for desired experimentation and data collection to occur.

The theory of the segregation model is that the process of how agent decisions aggregate can produce segregated communities in spite of the fact that none of the agents of the system have strong preferences for segregation and are even willing to be minorities within their local neighborhood. With this basic theory and a simple computational model, more complex questions can be examined in order to generate insights of relevance to intelligence analysts—insights that may be germane to the problem of community segregation itself, to epistemological and operational concerns associated with the collection of information on complex social systems like the society of interacting agents within a model, or as generator of future scenarios or counterfactuals for the purposes of hypothesis generation and evaluation that undergirds intelligence analysis.

As the remainder of this chapter demonstrates, relating model theory, experimental design, and the instantiation of the conceptual model itself in software is an

evolutionary process, constantly requiring tradeoffs. For example, computational speed and reliability may benefit from maintaining as much information as possible about the agents and their environment in a single location in software and memory. However, if experimental questions and designs require that each agent acts on its unique, perhaps misinformed or incomplete, view of the world, then it may be necessary for each agent to maintain their own individual version of reality in software, greatly increasing the memory and computational needs of the software. Such tradeoffs come in many forms and will often be discovered as model developers and users, analysts, and methodologists employ the model to address research questions and develop analytic products.

Understanding that models are developed iteratively according to their underlying theory, experimental goals, and other criteria, e.g. the availability of empirically relevant data, has significant consequences for how they fit into analytic tradecraft and organizations. Rather than treat models as finished software tools and products that analysts use to get particular results, in much the same way a calculator would be employed to solve an equation, computational models should be treated as living and evolving experimental platforms. Computational models, such as ABM, must constantly be adapted to suit changing conceptual concerns and technological circumstances. The addition of new theoretical entities within the models, such as new agent attributes or classes of behavior may be added, while changes in technology may demand models be reimplemented in order to capitalize on new hardware or software. Likewise, new experimental designs and analytic questions may demand that previous tradeoffs be reconsidered, often requiring major modifications to the code base, or the invention of

new model parameters and analytic metrics. Importantly, the learning process of greatest value to intelligence analysis comes from active participation in these processes, which provides the greatest insights into the features of the problem being investigated and the basis by which model development and experimental design can justify analytic findings by contextualizing output data into the larger research context.

Using a segregation model as a research tool requires developing a simple version of the model that can be extended in several different ways depending on the theoretical questions of analysts, concerns of policy makers, required experimental designs, and availability and compatibility with available data. The core model employed in this chapter was developed in NetLogo, and will be discussed in technical detail below in order to provide contrast to the significantly modified model structures required for performing particular experiments.⁴²²

Core Model Interface

The first portion of the core model is the interface. This is the graphical portion of the model that users interact with in order to setup, run, and interpret simulations. The simulation interface for the core segregation model is sparse, consisting of two buttons, two slider bars, a worldview, two graphs, and five variable monitors. The interface is depicted in Figure 5-6 below.

⁴²² Uri Wilensky, *NetLogo*, Center for Connected Learning and Computer-Based Modeling (Evanston, IL: Northwestern University, 1999), <http://ccl.northwestern.edu/netlogo/> (accessed on December 5, 2012)

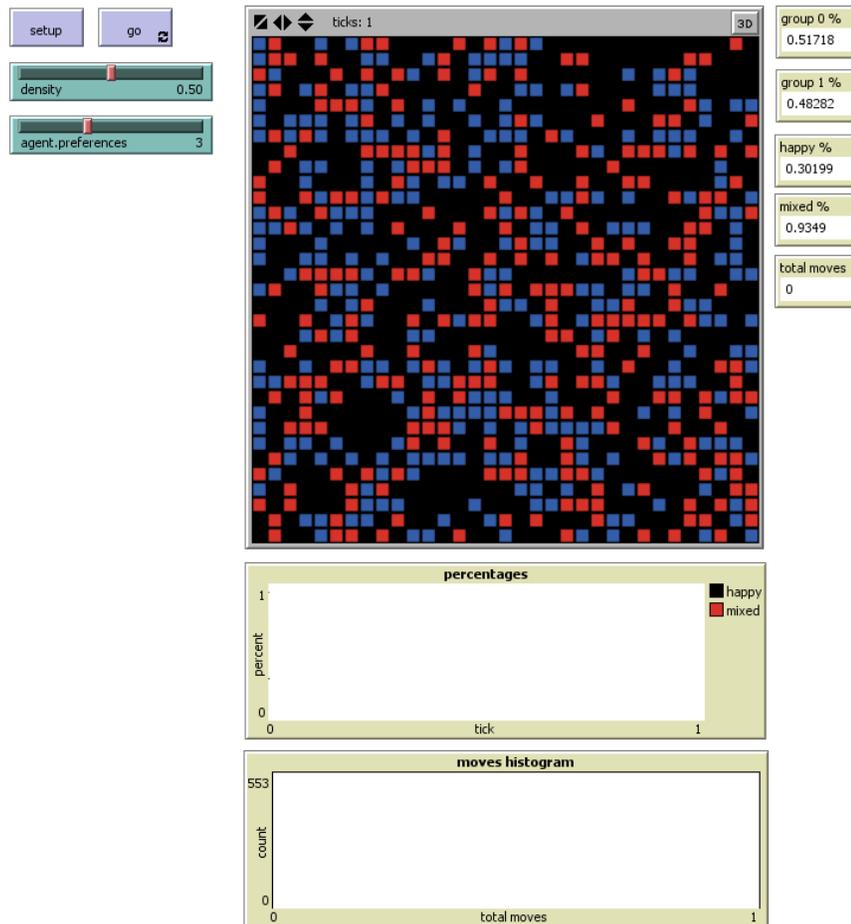


Figure 5-6: The Interface of the Core Segregation Model. The interface consists of four major sections for interacting with model users. On the upper left of the interface are two buttons in blue, and two sliders in green. The large window in the center of the interface is the worldview that depicts agents and the environment. In this case, agents appear as red or blue squares denoting their identities. Beneath the worldview are two graphs that update as the simulation runs, tracking statistics about the agent population by updating when instructed to do so. On the far right are several monitors that track population statistics numerically and update in real time.

The interface consists of four major sections for interacting with model users. On the upper left of the interface are two buttons in blue, and two sliders in green. The blue “setup” button is pressed to setup an agent population according to the parameters set on

the green sliders below. The blue “go” button is pressed to start the simulation and have it run until its stopping criterion is reached. The green slider bars control the population density, i.e. the number of agents in the population, and the agent preferences, which determine the number of neighbors of a common color each agent wants in order to be happy in their current location.

Population density is determined probabilistically by asking each patch on the landscape to draw a random number between zero and one, and if that number is less than the value of the “density” slider, a new agent is placed on that patch. The result is that the number of agents created for the simulation approximates the desired population density determined by the slider. Likewise, when agents are created, their individual preference is set to the value in the “agent.preferences” slider.

The large window in the center of the interface is the worldview that depicts agents and the environment. In this case, agents appear as red or blue squares denoting their identities. This window provides the primary window into the artificial society of the agents, showing where they reside at any moment in time. The worldview also shows the simulation’s time-step or “ticks” that have passed in the top of the window.

Beneath the worldview are two graphs that update as the simulation runs, tracking statistics about the agent population by updating when instructed to do so. Graphs are most frequently instructed to update based on model ticks, plotting new information whenever time advances. The graphs shown in the core model track the percent of agents that are happy and residing in mixed neighborhoods each tick, as well as a histogram that shows the distribution of the number of times each agent has moved to new locations on

the landscape. Tracking distributional information about agents is often an overlooked but significant benefit of ABM, enabled by its methodologically individualist design, because processes of social interaction and evolution often produce highly skewed distributions that do not conform to analysts' intuition.

On the far right of the interface are several monitors that track population statistics numerically and update in real time. While the information depicted in these monitors is often duplicated graphically, monitors are often employed for the purposes of debugging model code by ensuring specific variables are within their expected ranges or values, e.g. monitoring percentages to see if they ever achieve unallowable values that are less than 0 or greater than 100. Monitors are also employed to provide additional detail or resolution to graphs, such as in cases where the corresponding numerical values of plotted information cannot be easily or precisely read from the graph, or to track information that may not need to be displayed graphically but analysts may nevertheless wish to observe. Monitors on the interface display the percent of agents assigned to each group, nominally labeled group 0 and group 1 and colored as blue or red. Monitors also display the percent of agents that are happy, and the percent that are in mixed neighborhoods, as well as the total moves made by all agents in the population. Importantly, monitors employ a separate random number stream than the central one employed by the simulation itself, an issue that will be shown to be of great importance later.

Core Model Code

The use of the “.” in procedure and variable names is employed as a convention to keep NetLogo’s syntax consistent with other programming languages such as Java.

Unlike other programming languages, the “.” has no technical meaning in NetLogo but its use assists in making the code more familiar to readers accustomed to object-oriented programming languages.

Because variable names may often be quite common, five conventions are employed in order to differentiate variable types.⁴²³ These conventions are:

- “my.” is used as a prefix to identify variables that are owned by individual agents;
- “p.” is used as a prefix to identify variables that are owned by patches;
- “g.” is used as a prefix to identify variables that are global and always accessible;
- “_” is used as a prefix to identify a temporary variable that is not stored in memory once a procedure is completed;
- “?” is used as a suffix that denotes when a variable holds the value of TRUE or FALSE (T/F).

⁴²³ These coding conventions were developed by LMI’s Christopher Johnson and George Hart in 2010 as part of a modeling NetLogo-based modeling project.

```

globals
[
  ; no global variables are defined
]

patches-own
[
  ; patch variables are defined
]

```

The term “breed” declares an agent type by its plural and singular name, between opening and closing brackets. Because the words “agents” and “agent” are reserved in NetLogo and cannot be used as a name, the agents in the segregation model are named “schelling.agents” in plural and “schelling.agent” in the singular.

```

breed
[
  schelling.agents
  schelling.agent
]

schelling.agents-own
[
  my.id
  my.group
  my.preference
  my.moves
  my.happy?
  my.mixed?
]

```

In the core model, agents possess two classes of variables. The first class of variables is static and unchanging. These may be features that play no role in the model’s behavior, but are helpful for data analysis and assessment, such as assigning individual ids to each agent for the purpose of tracking them over time or in data outputs. Other static features include the agent’s group identity and neighborhood preference, which do not change once initialized in the core model. These variables are all identified above as “my.id”, “my.group”, and “my.preference”.

The second class of variables describes the agent's state and history. These variables may change over time and characterize the agent's unique experience within the system and its current condition or context that will motivate its future actions. In the core model, these dynamic variables include whether the agent is happy, in a mixed neighborhood, and the total number of times it has moved locations in its effort to satisfy its preference. These variables are identified by "my.happy?", "my.mixed?", and "my.moves" above.

Setting up the segregation model in code occurs when the procedure "setup" is called by pressing the "setup" button on the interface. The "setup" procedure initializes the model by going through the following steps:

- 1) clear all memory and reset the simulation time keeper;
- 2) set the default shape of all schelling.agents to be a square;
- 3) create a temporary variable "_id" and set it to zero; this will be used to assign individual id numbers to each schelling.agent when they are created;
- 4) ask each patch to consider creating a new schelling.agent based on a random draw between [0 1]; if the draw is \leq the value of "density" (located on the interface as a slider), the patch will "sprout" 1 new schelling.agent, otherwise it will remain empty;
- 5) for each schelling.agent being created:
 - a) set "my.id" to the value of "_id" and then advance "_id" by one for the next agent;
 - b) set "my.group" to a random number, 0 or 1;

- c) set its color to red or blue based on whether the group is 0 or 1;
 - d) set “my.preference” to the value of the preference slider on the interface;
 - e) set “my.happy?” to false;
 - f) set “my.mixed?” to false;
 - g) set “my.moves” counter to 0;
- 6) update the schelling.agents by having them see if they are happy and in mixed neighborhoods;
- 7) advance the simulation forward one time-step or “tick”;
- 8) update the plots on the simulation interface.

```

to setup
  __clear-all-and-reset-ticks
  set-default-shape schelling.agents "square"

  let _id 0
  ask patches
  [
    if random-float 1.0 < density
      [
        sprout-schelling.agents 1
        [
          set my.id _id
          set _id _id + 1
          set my.group random 2
          ifelse my.group = 0
            [set color blue]
            [set color red]
          set my.preference agent.preferences
          set my.happy? false
          set my.mixed? false
          set my.moves 0
        ]
      ]
  ]

  update.schelling.agents
  tick
  update.plots
end

```

Time proceeds in the simulation by calling the “go” procedure. This first begins by pressing the “go” button on the interface, and is repeated until the model’s stopping criterion is reached, which occurs when all schelling.agents are happy and have moved into neighborhoods that satisfy their preferences. Each time-step or “tick” consists of four activities:

- 1) Have all schelling.agents examine their neighborhoods and determine if they are happy and in a mixed neighborhood;
- 2) Have all unhappy schelling.agents move to new locations chosen at random;
- 3) Update all plots on the interface;

4) Advance the simulation forward one “tick”.

A final activity that occurs each time-step is to compare the state of the population of `schelling.agents` against the stopping criteria of the simulation. The simulation stops once there are no unhappy `schelling.agents`, i.e. when there are not any `schelling.agents` whose “`my.happy?`” is set to false.

```
to go
  update.schelling.agents
  move.schelling.agents
  tick
  update.plots
  if not any? schelling.agents with [not my.happy?]
    [stop]
end
```

Movement in the core model occurs when the procedure “`move.schelling.agents`” is called and instructs `schelling.agents` that are not happy with the current neighborhood to move to a new unoccupied location chosen at random. If a `schelling.agent` moves, its “`my.moves`” counter is incremented by one step. Both the random order in which `schelling.agents` move, and the random locations they move to, are determined by the simulation’s central random number stream.

```

to move.schelling.agents
  ask schelling.agents with [not my.happy?]
  [
    move-to one-of patches with [not any? schelling.agents-
      here]
    set my.moves my.moves + 1
  ]
end

```

Updating the state variables of the `schelling.agents`, specifically determining when they are happy and if they reside in mixed neighborhoods, occurs by calling the “`update.schelling.agents`” procedure. This procedure asks each `schelling.agent` to determine whether it is happy and in a mixed neighborhood. Happiness is determined by comparing the number of neighboring `schelling.agents` of the same identity with its “`my.preference`” variable. If the count of common neighbors is greater than or equal to “`my.preference`” the `schelling.agent` is happy, otherwise it is unhappy. Mixed neighborhoods are determined by the presence of any `schelling.agents` with identities from other groups as neighbors. The `schelling.agents` update their states in a random order determined by the simulation’s central random number stream.

```

to update.schelling.agents
  ask schelling.agents
  [
    set my.happy? count (schelling.agents-on neighbors) with
      [my.group = [my.group] of myself] >= my.preference set
    my.mixed? any? (schelling.agents-on neighbors) with
      [my.group != [my.group] of myself]
  ]
end

```

Plotting information in the core model occurs when the procedure “update.plots” is called, which then calls two subordinate procedures, each of which update different plots on the interface:

- 5) “update.percentages.plot” updates the line plot showing the percentage of schelling.agents that are happy and in mixed neighborhoods each tick;
- 6) “update.moves.histogram.plot” updates the histogram of the “my.moves” variable for all schelling.agents.

```
to update.plots
  update.percentages.plot
  update.moves.histogram.plot
end
```

The procedure “update.percentages.plot” performs the following steps:

- 7) set the current plot to the plot named “percentages” on the simulation interface;
- 8) set the current plot pen to “happy”, which tells the simulation to draw the line that tracks the percentage of schelling.agents that are happy this tick;
- 9) plot the computed percent happy variable, based on counting the total number of schelling.agents with their “my.happy?” variable equal to TRUE, divided by the total number schelling.agents in the system;
- 10) set the current plot pen to “mixed”, which tells the simulation to draw the line that tracks the percentage of schelling.agents that are in mixed neighborhoods this tick;

- 11) plot the computed percent in mixed neighborhoods variable, based on counting the total number of schelling.agents with their “my.happy?” variable equal to TRUE, divided by the total number schelling.agents in the system.

```
to update.percentages.plot
  set-current-plot "percentages"

  set-current-plot-pen "happy"
  plot count schelling.agents with [my.happy?] / count schelling.agents

  set-current-plot-pen "mixed"
  plot count schelling.agents with [my.mixed?] / count
    schelling.agents
end
```

Finally, the “update.moves.histogram.plot” procedure performs the following steps:

- 12) set the current plot to the plot named “moves histogram” on the simulation interface;
- 13) set the current range of the x-axis of the plot from 0 to the maximum value of all schelling.agents’ my.moves variable + 1:
 - a) adding 1 to the maximum value ensures a legal range in the case that all schelling.agents have moved zero times, which is the case during the initialization of the simulation;
- 14) set the current range of the y-axis of the plot from 0 to 1; this will be rescaled once the histogram is calculated, but it ensures that the y-axis will be rescaled dynamically as the counts for each my.moves value decrease as the agents settle at different rates;

15) set the number of bars employed in the simulation to be equal to the maximum value of my.moves among the population of schelling.agents + 1:

a) adding 1 to the maximum number of bars always exceeds zero, and means that each bar corresponds to an achievable, discrete number of moves by schelling.agents, i.e. the movement data is never binned into fractional units;

16) plot the histogram of the my.moves variable for the schelling.agent population.

```
to update.moves.histogram.plot
  set-current-plot "moves histogram"
  set-plot-x-range 0 (max [my.moves] of schelling.agents) + 1
  set-plot-y-range 0 1
  set-histogram-num-bars (max [my.moves] of schelling.agents) + 1
  histogram [my.moves] of schelling.agents
end
```

Interpreting the Core Model

Because models are abstractions, understanding precisely what they represent requires considering how the combination of their theory and technical design interact. In the core segregation model, agents are understood to be individuals or households that are seeking to satisfy their preferences with respect to their identity in a heterogeneous population. In the most basic form of the model, identity is an abstract notion that is generally interpreted to constitute race, ethnicity, religion, or other groupings.

Choosing a level of analysis upon which to interpret the meaning of identities affects the interpretation of what it means for an agent to reside in a neighborhood or

move locations on the landscape. For example, Schelling's original specification of the model interpreted identity as race, and location and movement as physical geography in order to examine the real-world emergence of segregated communities as a spatial phenomenon. However, interpreting identities as political party affiliation or membership in academic disciplines would imply a more abstract sense of location and movement, such as a willingness to offer bipartisan support to proposed legislation or participate on interdisciplinary research projects.

Model interpretation affects theory based on how the agents behave and interact with one another and their environment. Certain decision-making processes may be credible when interpreted within one framework, while inappropriate for others. Often model parameters may lack clear meaning, and must be redefined, parameterized, or invented in order support alternative interpretations of the model.

The extensions of the segregation model that follow maintain Schelling's original interpretive framework, where agents constitute individuals in an abstract, featureless environment. Although there is often a desire to place stylized and abstract agents into highly realistic environments, particularly as spatial data and Geographic Information Systems (GIS) become more available, the results of these efforts may prove misleading unless the agents are appropriately designed to perceive, move, and interact in realistic environments appropriately. For example, agents in the segregation model interpret their immediate environment as their Moore neighbors on a two-dimensional, featureless grid and move at random. Replacing the grid with a city or country map, or three-dimensional landscape would fundamentally require altering their behavioral rules and information

collection and processing requirements. For example, should agents in an apartment building be aware of and consider those who live above or beneath them in the same building? If agents appear to reside on a real landscape, is it appropriately dense so that the physical size of the agents relative to the environment's features is properly scaled and interactive?

Extending the Core Segregation Model

Given this study's theoretical and methodological orientation, increasing the realism of the segregation model is not warranted. Because of the model's regular appearance in ABM research, several others have employed it and extended it to address the issues noted above.⁴²⁴ While such extensions have demonstrated the potential of ABM-based research in the social sciences, particularly its ability to employ empirical information in the design of model environments and agent attributes, these efforts are of secondary importance in this study's examination of ABM in intelligence analysis.

Chapters 6, 7, and 8 that follow demonstrate different ways in which Schelling's segregation model can be employed in intelligence analysis. Each of these demonstrations addresses particular challenges associated with analytic tradecraft, such as theory development, coordinating collection and analysis, and tailored examinations of

⁴²⁴ Jason Barr and Troy Tassier, *Segregation and Strategic Neighborhood Interaction*, Department of Economics, Rutgers University, February 2007, <http://andromeda.rutgers.edu/~jmbarr/NYComp/BarrEEA.pdf> (accessed on November 8, 2012); Emin Dokumaci and William H. Sandholm, *Schelling Redux: An Evolutionary Dynamic Model of Residential Segregation*, Department of Economics, University of Wisconsin, June 19, 2007, <http://www.ssc.wisc.edu/~whs/research/schelling.pdf> (accessed on November 8, 2012); Abhinav Singh, Dmitri Vainchtein, and Howard Weiss, *Limit Sets for Natural Extensions of Schelling's Segregation Model*, June 15, 2010, <http://people.math.gatech.edu/~weiss/pub/Sch2June15.pdf> (accessed on November 8, 2012); and Nathan Berg, Ulrich Hoffrage, and Katarzyna Abramczuk, "Fast Acceptance by Common Experience: FACE-Recognition in Schelling's Model of Neighborhood Segregation," *Judgment and Decision Making*, Vol. 5, No. 5 (August 2010), pp. 391-410.

agency within complex adaptive systems. Together, they illuminate the general features and practices of a model-centric analysis and the unique opportunities and contributions of ABM as methodology.

Chapter 6: Developing Theories and Hypotheses

This chapter demonstrates how simple models can assist analysts in the development of theories about hypotheses that constitute the major intellectual thrusts behind the production of strategic intelligence. This is performed by modifying the segregation model discussed in Chapter 5, and examining how the use of coercion within mixed societies can alter settlement dynamics and patterns. The employment of the segregation model is fundamentally pedagogical and not intended to provide a comprehensive model of social dynamics that motivate contemporary security challenges such as sectarian relations and violence. Nevertheless, by showing how such a simple model may be extended, both theoretically and through experimental design, many of the problems intelligence analysts face can be addressed.

Coercion and Community Formation

The segregation model described a system in which all of the agents determined when and where to move voluntarily or of their own agency. However, social systems are not created by completely cooperative or voluntary actions, and are shaped by the explicit and implicit use of force. Force may be directed internally, within an identity group as a means for punishing those who defect from acceptable norms and traditions, or externally against rival out-groups in order to protect or extend the relative influence, power, and autonomy over others. Indeed, the process of state formation has been seen

as fundamentally coercive by anthropologists, sociologists, and historians alike, where armed groups exert force within their communities to gain control over the social and economic opportunities of the population, and then institutionalize their privileged positions through the development of political institutions and ideology.⁴²⁵

Enforcer Agents

In order to explore the ways in which coercion affects social dynamics in mixed societies the segregation model was modified by adding a new type of social agent.

These agents were called enforcers, and could focus their coercive power against other members of the same group, or agents from other groups. The introduction of enforcer agents required several new variables and agent logic, meaning that it was not simply a technical modification, but also required additional theorization and design choices.

Because the segregation model is a spatial model, the behaviors of agents are expressed through their movement rules and logic. In the original model, agents determined where to move voluntarily and at random based on the composition of their local neighborhoods. The introduction of enforcers alters this behavior by scanning their own neighborhoods and compelling targeted agents to move regardless of their happiness.

⁴²⁵ Robert Carneiro, "A Theory of the Origin of the State," *Science*, No. 169 (1970), pp. 733-738; Kent V. Flannery, "The Cultural Evolution of Civilizations," *Annual Review of Ecology and Systematics*, Vol. 3 (1972), pp. 399-426; William H. McNeill, *The Pursuit of Power: Technology, Armed Force, and Society Since A.D. 1000* (Chicago, IL: University of Chicago Press, 1982); Charles Tilly, *Coercion, Capital and European States: AD 990-1992* (Malden, MA: Blackwell Publishing, 1992); Jean Guilaine and Jean Zammit, *The Origins of War: Violence in Prehistory* (Malden, MA, Blackwell Publishing, 2005); Azar Gat, *War and Human Civilization* (New York, NY: Oxford University Press, 2008); and Michael Howard, *War and European History* (New York, NY: Oxford University Press, 2009).

From a theoretical perspective, the presence of enforcers implicitly challenges the basis of Schelling's original question about segregation. Because Schelling's initial conjecture was whether agents without strong in-group preferences could nevertheless develop segregated communities, inserting enforcers into the model admits the existence of agents with the very preferences Schelling sought to eliminate from the system, and endowed them with the capability to impose their will on others. As a result, enforcers should not be expected to fundamentally alter the qualitative findings of the original model—the emergence of segregated communities within a mixed society of agents. What is less clear, however, is determining the extent to which small numbers of enforcers might quantitatively affect the behavior of the system, whether focusing on in-group solidarity or out-group domination produces different dynamics, or what ancillary consequences might be caused by the presence of enforcers that are unattainable in its original specification.

From an operational perspective, introducing enforcers requires adding new features to the model. First, as indicated above, enforcers are likely to possess strong group preferences that are more powerful than those of ordinary agents. Whereas ordinary agents are assumed to be tolerant and willing to be minorities in their local neighborhoods (at least under the conditions that interested Schelling), enforcers would not be willing to be local minorities and possess much higher preferences with respect to the composition of their eight neighbors. As a result, a simple modification is simply to provide enforcers with a unique and more strenuous neighborhood preference that would prevent them from being content as a local minority.

The selection of enforcer preferences, however, has unexpected effects. As long as enforcers compel others to move, they will always generate homogeneous neighborhoods by driving away members of targeted groups. For example, an enforcer focused on in-group discipline would force members of its own group to move if they were in mixed neighborhoods. Thus, the enforcer's own neighborhood composition would drive it into minority status by reducing the presence of in-group neighbors. Likewise, an enforcer focused on out-group coercion would drive away members of the opposite group whenever possible, and therefore establish majority presence within its own local neighborhood. Therefore, the preference of an enforcer agent with respect to its neighborhood cannot be interpreted in the same way as non-enforcer agents, thus requiring a new theoretical interpretation of the preference parameter.

A new interpretation of the enforcer neighborhood preference parameter is the agent's individual aggressiveness, or the extent to which it seeks out opportunities for conflict. Although the need to reinterpret the preferences of enforcers as their aggressiveness is not immediately evident, it is consistent with the model's technical design and agent behavioral rules, as well as existing social theory. Within the model, enforcers move based on the same criteria as social agents, but because they drive their neighborhoods to unmixed states, preferences are one-sided and account for the number of agents with a shared identity in their immediate local environment. This means that an enforcer with a neighborhood preference of three will cease moving if it has a minimum of three of its eight neighbors of the same group, and it will continue to move to new locations if its preference is five and it has fewer than five common neighbors. Because

each move presents an opportunity to enter into a mixed community, this can be interpreted as a conflict-seeking or aggressive behavior. Under these circumstances, the higher the preference, the more the enforcer will move and therefore the more aggressive its behavior.

From the perspective of social theory, particularly the history of terrorism, the aggressiveness or willingness of terrorist groups to take risks has been an important variable.⁴²⁶ Moreover, the notion of reducing the aggressiveness of individuals and groups by embedding them in social communities with social ties that would restrain them is seen as an important way of reducing conflict and demobilizing terrorists, as demonstrated by the Palestinian Liberation Organization's (PLO's) handing of the Black September organization. These highly aggressive members were married off by the PLO's leadership in an effort to bind them to their communities and diminish their willingness to seek out new conflicts and potentially provoke strategically detrimental attacks.⁴²⁷

A second addition to the model is the inclusion of a vision radius for enforcers. This radius identifies the range with which enforcers exert their influence on their environment. As implemented, targeted non-enforcer agents within the vision radius of enforcers are compelled to move to new locations. In cases where enforcers operate against out-group agents, any agent within the radius of the opposite color of the enforcer is forced to move to a new location. In cases where enforcers operate against their own

⁴²⁶ Audry Kurth Cronin, *How Terrorism Ends: Understanding the Decline and Demise of Terrorist Campaigns* (Princeton, NJ: Princeton University Press, 2009).

⁴²⁷ Bruce Hoffman, "All You Need is Love: How the Terrorists Stopped Terrorism," *The Atlantic Monthly*, Vol. 288, No. 5 (December 2001), pp. 34-37.

group by enforcing in-group separation between identity groups, agents of the same color are forced to move to new locations if they are in mixed neighborhoods. Two enforcer agents within the segregation model are shown in Figure 6-1 below.

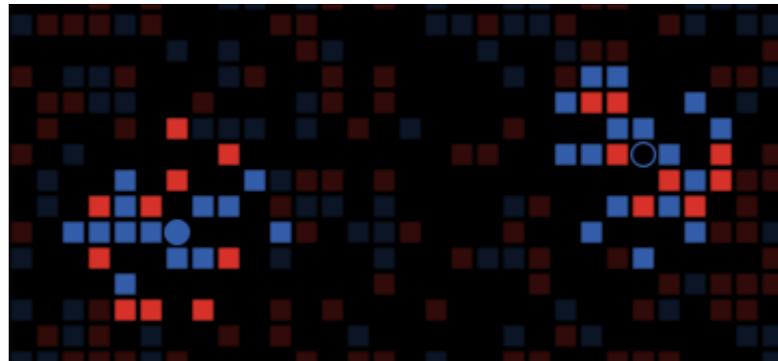


Figure 6-1: Depiction of Enforcer Agents. The circles on the left and right of the image above each show enforcer agents. The solid circle on the left shows an out-group enforcer, which in this case would compel all non-enforcer red agents (red squares) to move to new locations. The hollowed blue circle on the right shows an in-group agent, which would force all blue non-enforcer agents (blue squares) to move to new locations if they have any red neighbors.

For visualization purposes, all agents in Figure 6-1 have been darkened with the exception of enforcers, shown as colored blue circles, and those within their enforcement radius, shown as bright solid squares. On the left of the figure, a blue enforcer depicted as a solid blue circle targets out-group agents, i.e. red squares in its enforcement radius. On the right of the figure, a blue enforcer agent that is focused on in-group solidarity, depicted as a hollow blue circle, is shown targeting bright blue squares in its enforcement radius.

Several experiments were conducted employing enforcer agents in different forms. These experiments are discussed below. While their analysis is not comprehensive, two significant findings emerged. First, very small numbers of enforcers could have disproportionate quantitative effects on the overall spatial pattern of agents. While neighborhoods remained segregated, there were clear distinctions between those simulations with enforcers and those without. Second, the consequences of in-group vs. out-group enforcement were asymmetric. While resulting in a small difference in behavioral rules and algorithmic implementation, the consequences on settlement dynamics were extreme. An out-group orientation by enforcers largely extended the system's natural tendency to segregate, producing occasional waves of neighborhood destruction whenever enforcers entered into mixed neighborhoods. However, movement patterns became increasingly fragile and chaotic when enforcers focused on in-group solidarity and compelled members of their own group to move to new locations whenever an agent of the opposite group appeared. The reason for this is that the power to determine movement was indirectly and subtly transferred from small numbers of enforcers onto the entire population of agents from the alternative group: The appearance of a single agent with an out-group identity forced entire neighborhoods to be abandoned at once.

The results discussed below are based on a series of fifty distinct computational experiments where each combination of parameters was run fifty times, thus summarizing a total of 2,500 simulation runs. In each case, several model parameters were held constant, meaning that only three parameters varied across each simulation: the

number of enforcers in the population of red agents, the number of enforcers in the population of blue agents, and whether enforcers focused their coercive powers internally or externally with respect to their identity groups. In all other respects, the model was configured in an identical fashion for each simulation run, i.e. the number of agents in each group, their preferences, the size of the world, and the initial position of each agent were invariant across all simulations within and between experiments.

Out-Group Experiments

The parameters employed for Experiment 1, experiments with enforcers focused on out-group coercion, are shown in Table 6-1 below:

Table 6-1: Out-Group Enforcer Parameters.

Experiment Number	Total Red Agents	Total Blue Agents	Agent Neighborhood Preference	Red Enforcers	Blue Enforcers	Enforcer Neighborhood Preference	Enforcer Targets	Enforcer Range	World Size
1-1	750	750	3	0	0	5	Out-Group	2	50 X 50
1-2	750	750	3	3	0	5	Out-Group	2	50 X 50
1-3	750	750	3	7	0	5	Out-Group	2	50 X 50
1-4	750	750	3	11	0	5	Out-Group	2	50 X 50
1-5	750	750	3	15	0	5	Out-Group	2	50 X 50
1-6	750	750	3	0	3	5	Out-Group	2	50 X 50
1-7	750	750	3	3	3	5	Out-Group	2	50 X 50
1-8	750	750	3	7	3	5	Out-Group	2	50 X 50
1-9	750	750	3	11	3	5	Out-Group	2	50 X 50
1-10	750	750	3	15	3	5	Out-Group	2	50 X 50
1-11	750	750	3	0	7	5	Out-Group	2	50 X 50
1-12	750	750	3	3	7	5	Out-Group	2	50 X 50
1-13	750	750	3	7	7	5	Out-Group	2	50 X 50
1-14	750	750	3	11	7	5	Out-Group	2	50 X 50
1-15	750	750	3	15	7	5	Out-Group	2	50 X 50
1-16	750	750	3	0	11	5	Out-Group	2	50 X 50
1-17	750	750	3	3	11	5	Out-Group	2	50 X 50
1-18	750	750	3	7	11	5	Out-Group	2	50 X 50
1-19	750	750	3	11	11	5	Out-Group	2	50 X 50
1-20	750	750	3	15	11	5	Out-Group	2	50 X 50
1-21	750	750	3	0	15	5	Out-Group	2	50 X 50
1-22	750	750	3	3	15	5	Out-Group	2	50 X 50
1-23	750	750	3	7	15	5	Out-Group	2	50 X 50
1-24	750	750	3	11	15	5	Out-Group	2	50 X 50
1-25	750	750	3	15	15	5	Out-Group	2	50 X 50

The experimental design shown in Table 6-1 reveals a simple parameter sweep, incrementing the number of enforcers within the red and blue populations over the values [0 3 7 11 15] until all combinations have been run 50 times each. By tracking two simple parameters, the consequences of coercive power within societies can be better understood. The first parameter is the amount of time steps, or ticks, each run of the simulation requires in order to achieve a distribution of agents that satisfies each individual. The second parameter is the final percentage of agents that have at least one neighbor of the opposite group.

and the extent to which these results varied: The more enforcers within the population, the more variability the results displayed. Table 6-2 depicts the mean, median, max, min, and standard deviation of each configuration of model performed in accordance with the experimental plan in Table 6-1.

Table 6-2: Total Time Statistics for Out-Group Enforcement Experiments.

		Mean Total Time				
Enforcers Present		No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers		32.660	50.460	60.820	74.240	80.080
3 Red Enforcers		50.520	66.540	88.200	108.700	125.020
7 Red Enforcers		67.960	89.180	120.320	142.140	169.580
11 Red Enforcers		76.800	97.220	174.120	181.900	217.360
15 Red Enforcers		82.680	124.920	189.720	224.700	296.900

		Median Total Time				
Enforcers Present		No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers		32.000	44.500	59.500	68.500	79.000
3 Red Enforcers		48.000	63.000	73.000	104.000	107.000
7 Red Enforcers		66.000	82.000	106.000	112.000	141.000
11 Red Enforcers		73.000	86.000	136.500	148.000	187.500
15 Red Enforcers		78.000	112.000	153.500	200.500	270.500

		Max Total Time				
Enforcers Present		No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers		48.000	125.000	135.000	166.000	130.000
3 Red Enforcers		133.000	159.000	262.000	243.000	324.000
7 Red Enforcers		168.000	200.000	283.000	453.000	529.000
11 Red Enforcers		134.000	221.000	641.000	613.000	501.000
15 Red Enforcers		157.000	390.000	574.000	610.000	819.000

		Min Total Time				
Enforcers Present		No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers		24.000	28.000	30.000	35.000	40.000
3 Red Enforcers		28.000	33.000	32.000	46.000	58.000
7 Red Enforcers		28.000	40.000	45.000	45.000	71.000
11 Red Enforcers		33.000	39.000	56.000	60.000	69.000
15 Red Enforcers		43.000	47.000	50.000	72.000	91.000

		Standard Deviation Total Time				
Enforcers Present		No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers		5.359	22.621	19.128	26.622	22.227
3 Red Enforcers		21.415	26.740	45.023	45.528	60.915
7 Red Enforcers		28.932	34.979	61.325	94.268	99.683
11 Red Enforcers		22.599	45.884	110.865	112.686	97.036
15 Red Enforcers		24.135	59.707	108.775	119.028	153.955

The results in Table 6-2 above show effects that are not surprising with respect to the inclusion of agents willing to employ coercive force within social systems. The time required for the system to settle increases as more and more enforcers enter into the system. What is more illuminating is that as the number of enforcer agents in the system increase, the diversity of outcomes increases as well, producing results that are more extreme with respect to the time required for final settlement. This is seen in the increasing separation between the mean and the median, wider spreads between the min and max values attained in each experiment, and larger standard deviations that all correspond with larger numbers of enforcer agents in the system.

Percentage Mixed

A second metric examined across the experiments was the percentage of agents that resided in mixed neighborhoods at the conclusion of each simulation run. Once again, the results were consistent with the expected effects of coercive agents in social systems: Increased numbers of enforcers corresponded to decreasing percentages of agents residing in mixed neighborhoods. The results of each experiment with out-group coercive agents are shown in Figure 6-3.

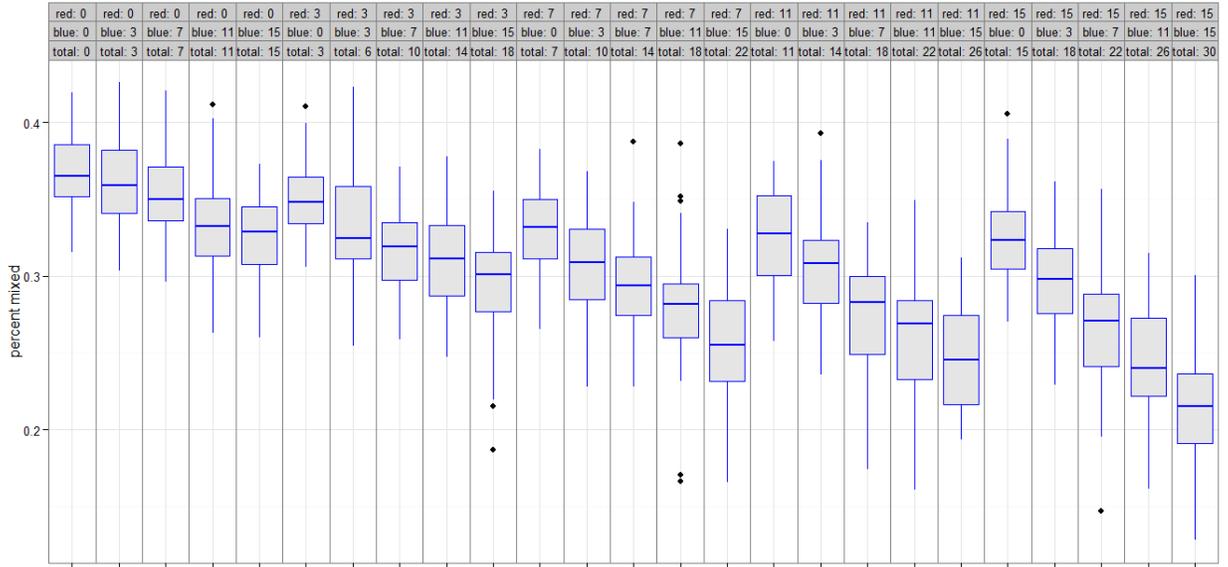


Figure 6-3: Final Percentage Mixed for Out-Group Enforcement Simulations. Each panel depicts the count of red enforcers, blue enforcers, and total enforcers, each of which focused on out-group coercion. The boxes provide a standard box-and-whiskers plot of the distribution of the percentage of agents residing in mixed neighborhoods at the end of each simulation.

The results of these experiments are further elaborated in Table 6-3 below. These results show a decreasing tendency for the system to produce mixed neighborhoods.

Table 6-3 depicts the mean, median, max, min, and standard deviation of each configuration of model performed in accordance with the experimental plan in Table 6-1.

Table 6-3: Percentage Mixed Statistics for Out-Group Enforcement Experiments.

		Mean Percentage Mixed				
Enforcers Present		No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers		0.369	0.360	0.354	0.332	0.245
3 Red Enforcers		0.351	0.332	0.317	0.312	0.293
7 Red Enforcers		0.330	0.309	0.297	0.281	0.254
11 Red Enforcers		0.325	0.308	0.274	0.263	0.246
15 Red Enforcers		0.325	0.298	0.265	0.245	0.215

		Median Percentage Mixed				
Enforcers Present		No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers		0.365	0.359	0.350	0.333	0.240
3 Red Enforcers		0.348	0.324	0.319	0.311	0.301
7 Red Enforcers		0.332	0.309	0.294	0.282	0.255
11 Red Enforcers		0.327	0.308	0.283	0.269	0.246
15 Red Enforcers		0.323	0.298	0.271	0.240	0.215

		Max Percentage Mixed				
Enforcers Present		No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers		0.419	0.426	0.421	0.411	0.315
3 Red Enforcers		0.411	0.423	0.371	0.378	0.355
7 Red Enforcers		0.383	0.368	0.387	0.386	0.331
11 Red Enforcers		0.375	0.393	0.335	0.349	0.312
15 Red Enforcers		0.405	0.361	0.357	0.315	0.301

		Min Percentage Mixed				
Enforcers Present		No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers		0.315	0.303	0.296	0.263	0.161
3 Red Enforcers		0.306	0.255	0.259	0.247	0.187
7 Red Enforcers		0.265	0.228	0.228	0.166	0.165
11 Red Enforcers		0.257	0.236	0.174	0.161	0.193
15 Red Enforcers		0.270	0.229	0.147	0.161	0.128

		Standard Deviation Percentage Mixed				
Enforcers Present		No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers		0.023	0.028	0.026	0.032	0.035
3 Red Enforcers		0.024	0.035	0.027	0.033	0.036
7 Red Enforcers		0.032	0.032	0.031	0.039	0.038
11 Red Enforcers		0.032	0.035	0.038	0.038	0.034
15 Red Enforcers		0.029	0.031	0.040	0.035	0.036

The measures of central tendency over these experiments support the claim that increases in enforcers decrease the presence of mixed neighborhoods. This is evident in both the mean and median values of the distributions for each experiment. However, an interesting dynamic can be seen in the dispersion of the results in each experiment. Whereas the variability between runs increased with the number of enforcers, the largest dispersion of results occurred at values less than the maximum number of enforcers. This is most likely explained by the fact that as the number of enforcers increases, their effectiveness with respect to compelling the population to segregate increases, thus resulting in less variability. With a small number of enforcers, the rate at which they can be placated by locating in homogeneous communities that suit the preferences is highly important—therefore there exists a tipping point with respect to which the diversity of results, as measured by the standard deviation of the percentage of agents in mixed neighborhoods, occurs.

In-Group Enforcement

The same experimental designs were repeated with enforcers who focused their coercive powers against members in their own groups. These experiments are described in Table 6-4 below, and were identical in to those performed in Table 6-1 in all other respects.

Table 6-4: In-Group Enforcer Parameters.

Experiment Number	Total Red Agents	Total Blue Agents	Agent Neighborhood Preference	Red Enforcers	Blue Enforcers	Enforcer Neighborhood Preference	Enforcer Targets	Enforcer Range	World Size
2-1	750	750	3	0	0	5	In-Group	2	50 X 50
2-2	750	750	3	3	0	5	In-Group	2	50 X 50
2-3	750	750	3	7	0	5	In-Group	2	50 X 50
2-4	750	750	3	11	0	5	In-Group	2	50 X 50
2-5	750	750	3	15	0	5	In-Group	2	50 X 50
2-6	750	750	3	0	3	5	In-Group	2	50 X 50
2-7	750	750	3	3	3	5	In-Group	2	50 X 50
2-8	750	750	3	7	3	5	In-Group	2	50 X 50
2-9	750	750	3	11	3	5	In-Group	2	50 X 50
2-10	750	750	3	15	3	5	In-Group	2	50 X 50
2-11	750	750	3	0	7	5	In-Group	2	50 X 50
2-12	750	750	3	3	7	5	In-Group	2	50 X 50
2-13	750	750	3	7	7	5	In-Group	2	50 X 50
2-14	750	750	3	11	7	5	In-Group	2	50 X 50
2-15	750	750	3	15	7	5	In-Group	2	50 X 50
2-16	750	750	3	0	11	5	In-Group	2	50 X 50
2-17	750	750	3	3	11	5	In-Group	2	50 X 50
2-18	750	750	3	7	11	5	In-Group	2	50 X 50
2-19	750	750	3	11	11	5	In-Group	2	50 X 50
2-20	750	750	3	15	11	5	In-Group	2	50 X 50
2-21	750	750	3	0	15	5	In-Group	2	50 X 50
2-22	750	750	3	3	15	5	In-Group	2	50 X 50
2-23	750	750	3	7	15	5	In-Group	2	50 X 50
2-24	750	750	3	11	15	5	In-Group	2	50 X 50
2-25	750	750	3	15	15	5	In-Group	2	50 X 50

The experimental design reveals a simple parameter sweep, incrementing the number of enforcers within the red and blue populations over the values [0 3 7 11 15] until all combinations have been run 50 times each. The same metrics of time required to completion and percentage mixed are examined below.

Time Steps

The presence of in-group enforcers has a significantly different effect on the time required for the system to settle into a state where every agent was happy. In fact, the enforcer preferences and actions were so powerful, that the simulation needed to be terminated at 10,000 ticks if it could not find a state that that satisfied every agent. This alone is significant when compared with the maximum amount of time experienced with out-group enforcers under otherwise identical conditions: 819 ticks. This difference

suggests a qualitative change in the dynamics of coercion results when aggressors focus on members of their own group in an effort to instill discipline, preserve purity, or otherwise deny group members from interacting with outsiders. Figure 6-4 below shows the distribution of simulation times by experiment.

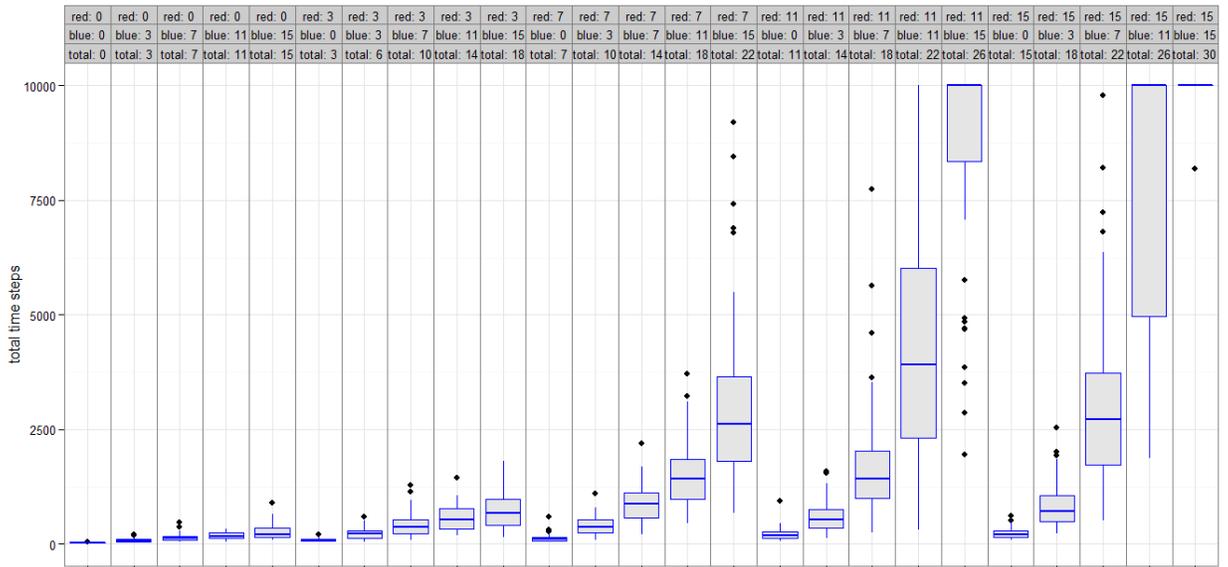


Figure 6-4: Total Time Steps for In-Group Enforcement Simulations. Each panel depicts the count of red enforcers, blue enforcers, and total enforcers, each of which focused on out-group coercion. The boxes provide a standard box-and-whiskers plot of the distribution of the time steps required for the simulation to reach a configuration where all agents were happy. Importantly, the simulation was terminated after 10,000 time steps if it had not yet discovered a stable spatial configuration.

Additional details of these experiments are shown in Table 6-5 below, where the measures of central tendency of each experiment with respect to time to settlement are shown.

Table 6-5: Total Time Statistics for In-Group Enforcement Experiments.

Mean Total Time

Enforcers Present	No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers	33.68	87.66	146.06	187.44	283.66
3 Red Enforcers	84.38	239.6	419.72	574.62	739.72
7 Red Enforcers	126.14	402.64	886.48	1538.6	3054.56
11 Red Enforcers	210.04	588.84	1756.56	4393.4	8683.26
15 Red Enforcers	242.86	865.62	3136.7	7801.7	9963.72

Median Total Time

Enforcers Present	No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers	33	74.5	135.5	176.5	208.5
3 Red Enforcers	73	230.5	373.5	536.5	683
7 Red Enforcers	109	373	879.5	1421.5	2617
11 Red Enforcers	185	525	1421	3905.5	10000
15 Red Enforcers	210.5	717.5	2710	10000	10000

Max Total Time

Enforcers Present	No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers	53	216	471	336	892
3 Red Enforcers	207	605	1281	1444	1805
7 Red Enforcers	599	1109	2192	3717	9196
11 Red Enforcers	940	1586	7743	10000	10000
15 Red Enforcers	609	2533	9775	10000	10000

Min Total Time

Enforcers Present	No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers	25	34	44	56	83
3 Red Enforcers	41	52	96	185	157
7 Red Enforcers	44	94	219	448	679
11 Red Enforcers	64	124	246	309	1945
15 Red Enforcers	88	237	519	1867	8186

Standard Deviation Total Time

Enforcers Present	No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers	5.72655825	44.03412869	83.17976844	76.45117423	185.1299169
3 Red Enforcers	33.89713732	131.1850228	252.0723932	285.8331817	420.352575
7 Red Enforcers	89.46211375	215.3861042	405.9339064	764.8678424	1934.313285
11 Red Enforcers	136.2615971	332.1457364	1356.651522	2695.925739	2315.236011
15 Red Enforcers	123.5440944	493.7409422	2007.361994	2779.521558	256.5383402

The results regarding time for the in-group experiments reveal that the larger the population of enforcers, the more difficult it becomes for the system to find a satisfactory distribution of agents in which each is happy with their neighborhoods. Importantly, in cases where the numbers of enforcers are the largest, the median outcomes are for the system not to settle, suggesting that the outliers are those rare cases in which a stable configuration is reached.

Percentage Mixed

The percentage of agents residing in mixed neighborhoods at the conclusion of simulation runs provides greater insights into the dynamics of coercion. The results of each simulation run by experiment are shown in Figure 6-5 below.

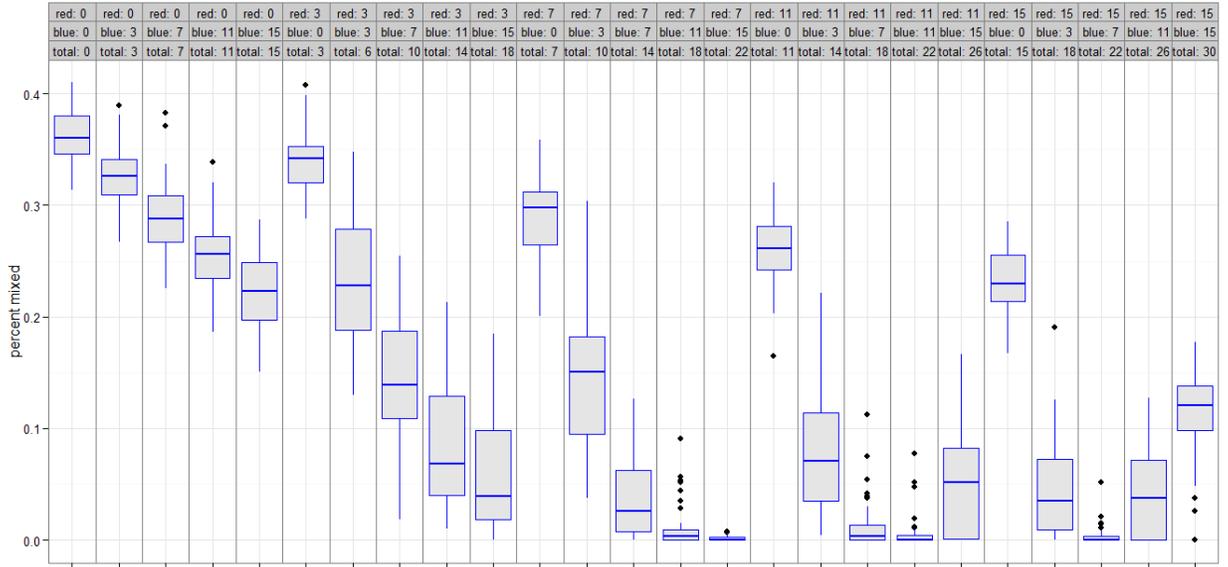


Figure 6-5: Final Percentage Mixed for In-Group Enforcement Simulations. Each panel depicts the count of red enforcers, blue enforcers, and total enforcers, each of which focused on in-group coercion. The boxes provide a standard box-and-whiskers plot of the distribution of the percentage of agents residing in mixed neighborhoods at the end of each simulation.

When examining the results of the experiments, it is important to consider the data on simulation time and the fact that simulations were terminated at 10,000 ticks if they had not reached an acceptable spatial configuration. As a result, the numbers of agents residing in mixed neighborhoods are artificially high, with the data reflecting the state of the simulation at its terminating point, and not at an acceptable configuration. The circumstances that can create artificially high, temporary mixed neighborhoods can be seen in Figure 6-6 below.

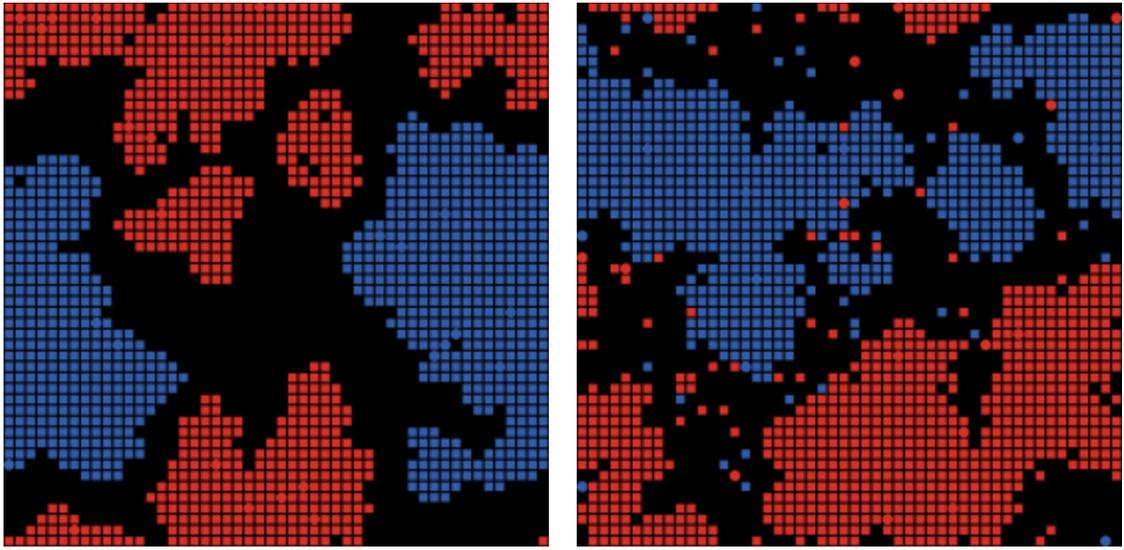


Figure 6-6: Sample Final Settlement Results of In-group Enforcement. The result of run number 1169 and 1173 as part of Experiment 2-24 described in Table 6-4. The figure on the left showed a settled landscape for run 1173, where the population was able to discover a solution where all agents were happy with their local circumstances. On the right, run 1169 is depicted, which shows the state of the population at tick 10,000 when the simulation terminated despite the fact that many agents were unhappy with their local neighborhoods.

These figures show the final landscapes of two simulations with high numbers of in-group enforcers. As the left image in Figure 6-6 shows, no mixed neighborhoods existed in a case where all agents are happy. However, the image on the right of Figure 6-6 shows a simulation that was unable to complete due to time, revealing the existence of several mixed neighborhoods formed by displaced agents and unhappy enforcers. This result is representative of many of the simulations with high levels of in-group enforcers, and is not seen in any of the comparable cases involving out-group enforcers.

The effects of in-group enforcement are clearer when the measures of central tendency are examined with respect to mixed neighborhoods. These are show in Table 6-6 below.

Table 6-6: Percentage Mixed Statistics for In-Group Enforcement Experiments.

Mean Percentage Mixed

Enforcers Present	No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers	0.3634	0.32672	0.290893333	0.25512	0.038906667
3 Red Enforcers	0.338946667	0.232813333	0.143813333	0.085653333	0.060106667
7 Red Enforcers	0.290213333	0.148253333	0.036586667	0.010413333	0.001373333
11 Red Enforcers	0.258706667	0.08144	0.012373333	0.00568	0.04832
15 Red Enforcers	0.232826667	0.045253333	0.003706667	0.038906667	0.11416

Median Percentage Mixed

Enforcers Present	No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers	0.360333333	0.326333333	0.287666667	0.256	0.037666667
3 Red Enforcers	0.342333333	0.227666667	0.138666667	0.068666667	0.039
7 Red Enforcers	0.297666667	0.151	0.026	0.003666667	0
11 Red Enforcers	0.261666667	0.070666667	0.003666667	0	0.052
15 Red Enforcers	0.229333333	0.035	0	0.037666667	0.120333333

Max Percentage Mixed

Enforcers Present	No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers	0.41	0.389333333	0.382666667	0.338666667	0.127333333
3 Red Enforcers	0.408	0.348	0.254666667	0.213333333	0.184666667
7 Red Enforcers	0.358666667	0.304	0.126666667	0.090666667	0.007333333
11 Red Enforcers	0.32	0.221333333	0.112666667	0.077333333	0.166666667
15 Red Enforcers	0.285333333	0.190666667	0.052	0.127333333	0.177333333

Min Percentage Mixed

Enforcers Present	No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers	0.314	0.267333333	0.225333333	0.186666667	0
3 Red Enforcers	0.288	0.13	0.018666667	0.01	0
7 Red Enforcers	0.200666667	0.037333333	0	0	0
11 Red Enforcers	0.164666667	0.004666667	0	0	0
15 Red Enforcers	0.167333333	0	0	0	0

Standard Deviation Percentage Mixed

Enforcers Present	No Blue Enforcers	3 Blue Enforcers	7 Blue Enforcers	11 Blue Enforcers	15 Blue Enforcers
No Red Enforcers	0.023609157	0.025638633	0.031302666	0.031895646	0.040159266
3 Red Enforcers	0.027144725	0.057169246	0.060355717	0.056425536	0.05476155
7 Red Enforcers	0.032727402	0.065845816	0.03611085	0.018565757	0.002053303
11 Red Enforcers	0.030499815	0.057534272	0.021628553	0.014503258	0.045072125
15 Red Enforcers	0.030073096	0.042240138	0.00857054	0.040159266	0.037215694

The measures of central tendency reveal how the presence of in-group enforcers prevents the system from achieving a stable distribution of agents. However, for the reasons noted earlier, the data shows a non-monotonic response with respect to the percentage of agents in mixed neighborhoods, an artifact of terminating simulations after 10,000 ticks if a satisfactory configuration could not be found beforehand. Therefore, the higher levels of mixing associated with increasing numbers of enforcers is the result of the configuration of agents still groping for locations in which they will be happy and free from the coercive powers of enforcers.

The most revealing metric associated with in-group experiments is the minimum percentage mixed found in each configuration of the model. As the number of enforcers increased, the system achieved the total segregation between groups, indicating the successful completion of the will of the enforcers. Importantly, this level of segregation was unachievable in comparable configurations of the model with out-group enforcers.

Summary

The introduction of enforcers into the segregation model does not constitute a comprehensive study or examination of coercion in society. However, it does provide a preliminary assessment of how coercive power may significantly enhance natural tendencies for systems to self-organize and segregate along the fissures between identity groups. Of greatest significance are the disproportionate effects that a small number of actors committed to the use of force can exert on the society as a whole. Recalling that the population of agents employed in the experiments shown in Tables 6-1 and 6-4 consisted of 800 individual agents—400 red and 400 blue—the fact that such major differences

could result from very small numbers of coercive actors is significant. With a maximum of 15 enforcers in each group, the settlement patterns within the population varied significantly based on how coercive force was employed. In both cases, the small numbers of enforcers relative to the entire population of agents revealed how actors whose preferences and behaviors are far from the norm of the larger population can disproportionately affect the system as a whole. This is important, because it demonstrates the need to consider the behaviors and actions of small groups and individuals that reside in the tails or larger distributions and whose presence, beliefs, and actions may not be captured by narratives and policies that focus on the average person.

From a substantive standing, the importance of in-group coercion begs new questions about the use of force and society. Violence between identity groups may create difficult political and social circumstances that must be managed, but as the model results suggest, socially acceptable configurations of society can be arrived at as long as the use of force remains confined to relatively small areas or numbers of actors. However, if agents use their coercive power to maintain internal discipline by targeting members of their own group, society itself may never settle into an acceptable configuration. Although real-world conflict may not manifest as endless spatial shuffling, the inability to settle into a stable configuration suggests many metaphorical interpretations that indicate the inability of coerced groups to enter into, or remain engaged in, larger social, political, and economic institutions without the risk of punishment from those committed to isolating their own group from the larger society.

An additional substantive finding concerns the magnitudes of the segregation effects achieved by in-group enforcers relative to out-group enforcers. The differences in outcomes suggests that smaller, weaker groups would be most successful at achieving their goals by terrorizing and disciplining their own identity groups, even if they would prefer to use force against out-group targets. Therefore, it is by extension likely that selection of targets by terrorist or insurgent groups may be an indicator of their own assessments of relative strength, and not simply a matter of ideology. The model results suggest that smaller groups could maximize their strategic influence by disciplining their own population groups rather than attacking members of rival groups. Moreover, because the relative effectiveness of creating homogeneous societies is higher by policing one's own group than in coercing other groups, the tendency for groups to commit inter-group violence may be interpreted as a belief in their strength, moral beliefs about the acceptable use of violence, and a willingness to accept less-efficient actions due to additional strategic considerations. Likewise, groups that shift the violence from out-group to in-group targets may be weakening, and lack the capability to effectively exert their will on rivals. In these cases, the restoration of internal discipline within the community may be viewed as necessary to demonstrating their strength and resolve, and achieving their goals of social segregation by employing fewer resources. Alternatively, groups that change from in-group to out-group targets may be growing in size, allowing them to achieve the same effects by focusing on socially and morally preferable targets, and extending the range of their concerns beyond military effectiveness to include factors such as political and social legitimacy.

The substantive conclusions regarding in-group and out-group violence and targeting should be regarded as speculative: They are hypotheses that are suggested by the model but are not proven by it. The model employed is highly abstract, and is best employed for aiding analysts by indicating new lines of research. The inferences from the model's behavior are not empirical findings, but they provide hints as to the kinds of information that might be collected and evaluated to test hypotheses suggested by the model. These may include monitoring the spatial interactions of groups in mixed societies, as well as other indicators of intergroup interactions, such as patterns of business investments, charitable giving and services, education and recreational activities, etc. Additional indicators may include the characteristics of identity-based political rhetoric, crime, violence, and the emergence of ethno-sectarian gangs or paramilitary organizations.

From a methodological perspective, the simple technical modifications to the segregation model performed by adding enforcer agents showed the utility and flexibility of ABM. The ease of adding new types of agents with unique behaviors encouraged experimentation and theory building. The need for introducing new parameters or reinterpreting existing ones, such as reconsidering how an agent's neighborhood preferences corresponded to aggressiveness in the special case of enforcer agents, revealed important linkages between methodology, technology, and theory, in which computational models are the intersection. While the simple segregation model originally developed by Schelling is overly simple to explore individual cases, by continually adding to its representation of social actors through new types, behaviors, and

attributes, and modifying the environment in which agents interact, the model can become increasingly realistic and specific, i.e. narrowed to represent specific cases of interest to policy makers and analysts. Thus, simpler models like the one employed here may be useful for generating new hypotheses, and then serve as vehicles aiding analysts develop, refine, and assess new theories and hypotheses by making models increasingly complex.

The methodological application of ABM, specifically the processes by which analysts can develop, analyze, adapt, and inference from computational models, can provide the centerpiece of a model-centric tradecraft. However, these activities are predominantly analytical and, while simple to perform within the artificial societies instantiated in ABMs, are not unique to them. These unique contributions of ABM are discussed in the following two chapters.

Chapter 7: Coordination between Collection and Analysis

Contemporary relations between intelligence analysts and collectors are increasingly important given the combination of new sources of information, means for collecting it, proliferation of potential types and sources of threats, and increased scrutiny with which intelligence agencies operate. However, relationships between analysts and collectors are also regarded as ad hoc and poorly systematized, remaining one of the perpetual challenges in the practice of intelligence.⁴²⁸ While the intelligence cycle, the most commonly used model of the intelligence process, places analysts in the position of evaluating intelligence information that has been collected according to the priorities of policy makers, this is rarely the way intelligence actually works.⁴²⁹ Instead, collection priorities are often developed by analysts, collectors, and managers working together in anticipation of consumers' future concerns.

The need for analysts and collectors to work together extends beyond matters of organizational or operational practicalities. Bruce, one of the most forceful proponents of improving relations between analysts and collectors, argued that analysts must understand the nuances of collection, and familiarize themselves with collection sources and methods in order to improve the quality of their assessments and guard against denial

⁴²⁸ Interview with Paul Pillar, Georgetown University, February 1, 2012.

⁴²⁹ Arthur S. Hulnick, "The Intelligence Cycle," in Loch K. Johnson and James J. Wirtz, eds., *Intelligence: The Secret World of Spies* (New York, NY: Oxford University Press, 2011), pp. 65-77.

and deception. Likewise, he argued that collectors benefit from analytic input with respect to improving targeting and recruiting, the vetting of sources, cross-validating information, and identifying intelligence gaps.⁴³⁰

On the collection side, the single most important factor accounting for failure is the impact of intelligence denial—namely, effective countermeasures taken by an intelligence target that prevented successful collection against it. The impact of denial, as nearly all these cases illustrate, is missing information needed for analysis. On the analytical side, the failure to correct for the impact of missing information, when combined with a lack of imagination, is an almost surefire predictor of analytical failure....

Better collection...requires greater analyst engagement and expertise; deeper understanding by analysts of the technical disciplines as well as the human recruitment and vetting process; and better guidance and direction in the requirements process where analysts have major, if often unfulfilled, responsibilities.⁴³¹

Furthering the argument regarding the need for analysts and collectors to coordinate, George noted the emergence of a new career path for analysts, called “targeting analysis,” which emphasizes the production of analysis to support intelligence collectors, rather than the traditional needs of policy consumers that focus on politics, leadership, economics, and military affairs.

Over the past nine years, the CIA has reverted to an organization that is more like a war-fighting agency than a twenty-first century knowledge-based think-tank. In a sense, it is back to the future. The CIA is now a re-born Office of Strategic Services (OSS). That 1940s model placed emphasis on wartime actions and operations, not analysis and foresight. Just as the OSS hired historians and economists to better fight Hitler, the NCS [National Clandestine Service] is harnessing more and more analytic talent to target terrorists and proliferators. These are short-term

⁴³⁰ James B. Bruce, “The Missing Link: the Analyst-Collector Relationship,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, pp. 191-210.

⁴³¹ James B. Bruce, “The Missing Link: the Analyst-Collector Relationship,” in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, pp. 191-192, and 203.

operational plans, which will have an immediate and important payoff. A “targeting analyst” discipline, separate from the traditional military, political, or economic analyst, is now a promising analytic career option. Working in the NCS-led Counter-terrorism Center (CTC) or its Counter-proliferation Division (CPD) are now more rewarded than researching trends or writing papers on African or Latin American countries. The “one Agency, one mission” mantra heard in Langley these days is a metaphor for having the NCS collection efforts and operational missions set priorities on allocating resources on analysis.⁴³²

Agent-Based Modeling and Intelligence Collection

ABM provides analysts and collectors with new opportunities to coordinate in ways that other formal models have been unable to provide. Because the centerpiece of ABM is the creation of an artificial society, rather than an equation, simulations maintain the system’s full dimensionality, allowing for simulated systems to be examined from multiple perspectives simultaneously. ABMs preserve a complete representation of the system as it moves through time, rather than compressing many independent variables into a single dependent one. As a result, ABMs provide analysts with the ability to simulate social systems in order to explore their properties, and also assess potential collection strategies and capabilities within it. This allows for analysts to better understand the ways in which collection efforts may be biased or misleading given specified social dynamics, prioritize collection targets based on analytic interests, understand vulnerabilities to adversary denial and deception, and more.

In addition to simulating alternative collection strategies against an intelligence target, employing evolutionary computation, machine learning, and other techniques may

⁴³² Roger Z. George, “Reflections on CIA Analysis: Is it Finished?” *Intelligence and National Security*, Vol. 26, No. 1 (2011), p. 77.

allow analysts and collectors to capitalize on computational technology in the design of collection systems. This contribution is significant because it enables machinery to assist in the search for solutions to intelligence collection problems, allowing for analysts and collectors to confront the problems of bounded rationality, satisficing, and limited search capabilities that undergird the problem of mindsets in analysis. Therefore the intersection between theoretical models of agent behavior and technologies for searching large possibilities of model configurations allows for formal models to assist in activities normally regarded as creative and beyond the reach of rigorous methodologies.

The remainder of this chapter is devoted to simulating the collection of intelligence information within the basic segregation model. This is done in two interrelated ways. First, three types of collection capabilities are simulated, each of which provides different a perspective on the agent population. Because the true state of the simulation's agent population is always known, potential biases in proposed collection systems can be uncovered, allowing ABM to aid collectors and analysts in tailoring strategies for particular targets, and understanding how to better interpret available sources and methods. Second, a genetic algorithm is employed to show how evolutionary computation can be employed to search for collection capabilities and strategies in ways that cannot be done without the assistance of computers. Together, the combination of social simulation, the simulation of intelligence collection, and the employment of computational techniques for searching across novel specifications of both allows ABM to provide the intelligence community with opportunities for

structuring relationships between analysts and collectors in ways that other formal models have not.

Collection Capabilities

Simulating alternative collection capabilities required several changes to the underlying technical structure of the core segregation model specified in Chapter 5. First and foremost, the simulation of collection capabilities introduced new kinds of agents into the simulation and demonstrated that agents within the model do not need to be social actors but can also be technical systems and processes as well. Within the software, these new agents were called “sensor.agents” and represented surveys, human agents, and technical sensors that each behaved differently as they gathered information on the movements, locations, and neighborhoods of social agents. For reasons of experimental design, a third class of agents called “sensor.portfolios” was created as well. These “sensor.portfolios” owned several sensors, often of different types, in order to fuse information and provide a more comprehensive perspective on the system derived from multiple sources and methods.

A second change was a modification to the model’s underlying technical structure and the management of the random number generators. This change was necessary in order to ensure that sensors could gather information on the population without affecting the behavior of social agents. By decoupling the behavior of agents from the operations of sensors, alternative collection capabilities and strategies could be compared against identical scenarios where agents behave identically. Through the simulation of three alternative types of sensors, analysts and collectors can better understand how surveys,

human agents, and technical sensors each have different information collection profiles that must be accounted for in assessments.

Surveys

The simulation of surveys replicated the collection of responses from a random sample of social agents to questions about their individual status and history. Surveys collected information on whether social agents resided in mixed neighborhoods, are currently happy, how many times they have moved, the last time they moved, and their current location. With this information from a sample of the population, the systemic features of the entire population can be estimated and compared with the true values of the population.

Agents respond to surveys honestly or dishonestly. When answering a survey honestly, each agent provides its id, group membership, total moves it has made, the last time step it moved, whether it is happy, and whether it is in a mixed neighborhood. If providing a dishonest answer, the agent's behavior is more complex due to the need to create a plausible set of responses that are obviously false. First, the social agent's true identification number is provided in their answers for the purposes of debugging and fusing information. Importantly, the sensor that receives this list has no access to the identification number variable and therefore cannot know if the agent being surveyed is responding honestly. Once completed, the agent provides a name at random, selecting one from other members of the population. Afterward, random values for group membership and location are provided.

With respect to movement values, these are selected at random within specified logical constraints. The total number of moves reported by the agent cannot exceed the total number of ticks that have passed in the simulation, and the total reported number of moves must be the sum of the reported voluntary and forced moves (if enforcers are present within the population); likewise, the last reported move must be consistent with the total number of moves, ensuring that an agent that reported 10 total moves does not assert its last time moved was tick 3. Once all questions regarding the agent's movement history have been answered, the agent randomly responds to questions about whether it is happy, in a mixed neighborhood, and its number of neighboring agents. After all the information has been gathered, the agent returns the list to the sensor for evaluation.

Surveys were characterized by four basic parameters:

- Sample size as a percentage of the population of social agents in the system;
- Reporting frequency as a probability that the survey gathers and reports information on a sample of the population on each simulation tick;
- Longitudinal sampling characterized as the probability that each social agent in a previous sample will be retained in a future sample;
- Reporting accuracy characterized as the probability that social agents respond honestly to survey takers.

An important feature of surveys in the context of the segregation model is that they collect information on the population as a random (or longitudinal) sample. They are not spatial in character and rely on information provided voluntarily by respondents.

Importantly, this differentiated surveys from the two other forms of information collection, human agents and technical sensors.

Human Agents

Human agents were simulated as members of the population who gathered information on their community and reported their observations back to intelligence analysts. Like surveys, they gathered information on whether agents resided in mixed neighborhoods, are currently happy, how many times they have moved, their current location, and their number of neighbors. Unlike surveys, however, human agents gathered information in a spatial context by observing the status of social agents in their own local communities. As a result, human agents are mobile and move within the population.

In the simulation, human agents, which are a subclass of “sensor.agents,” shadow social “schelling.agents.” Once a social agent is selected to be the human agent source which reports on its surroundings, the “sensor.agent” will follow it from tick to tick, and collect information from the vantage point of the social agent that is its human source. In this sense, the division between the social agents and the sensor agents within the model is a technical artifact of separating social agents from the sensors that gather information on them in the software itself. While this is not problematic in the case of surveys and technical sensors, real-world human sources are simultaneously collection assets and members of the target population. Therefore, a distinction exists between the technical implementation of sensors as operating independently in code and the conceptual notion of human sources as being embedded within the target population.

Human sensors always move to the location of their informing social agent prior to collecting information, ensuring that the spatial information they collect is based on the view of their source. If the human sensor is not located on the same patch, the “sensor.agent” moves to that location and updates the patches in view for collection, stored in the variable called “my.patches.in.view”. Once in place, human agents asked social agents within the radius “my.report.range” to provide information identical to that collected by surveys. This provided them with the same substantive information as surveys, but with a spatially based rather than random sample.

Human agents are characterized by three variables:

- Collection radius as a percentage of the maximum distance on the landscape that human agents observe and report on;
- Reporting frequency as a probability that a human agent gathers and reports information on the social agents within their collection radius in a given simulation tick;
- Reporting accuracy as a probability that the information reported on each social agent by the human source is accurate.

In addition to these three variables, a fourth variable determines which social agent serves as the human source, thus determining the movement and spatial trajectory of the human source over the simulation. This variable was set by drawing a random number between [0 1], scaling it against the population size of social agents, and then selecting the agent with the resulting id. For example, with 1500 social agents and a random draw value of

.3402862, the resulting value would assign the agent with “my.id” equals 510 as the human source.

Technical Agents

Technical sensors, referred to in code as “technical”, simulated the operations of collection sensors that can resolve the identity and location of individuals within a social system via technical means. Importantly, this does not represent a particular collection system, but a more abstracted notion of how technological capabilities might be employed to gather information on populations and what their relative strengths and weaknesses might be. As a result, whether the system in question is space-based imagery, networks of street cameras, the exploitation of personal communication technologies, or aerial surveillance systems such as small Unmanned Aerial Vehicles (UAVs) or large blimps, the basic analytic processes of aggregating data from technical sensors is the model’s focus.

Because technical sensors are largely stand-off and impersonal, many of the variables that might be gathered from surveys or human agents are not accessible. Simulated technical systems can only identify the group to which an agent belongs and their spatial location—these systems do not maintain an extended history of the individuals they observed or their movements, and therefore have no information on how many times an agent has moved or its social preferences, denying assessments of the social agent’s happiness with its current neighborhood. Moreover, because technical sensors cannot ask agents if they reside in mixed neighborhoods, they must infer whether an observed social agent is adjacent to any agents from other groups. By contrast,

surveys and human agents collecting information on the social population can gather information on the state and history of individual agents—all variables that are beyond the reach of technical sensors. Technical sensors are characterized by three parameters:

- The sensor's collection radius as a percentage of the maximum distance on the landscape;
- The sensor's reporting frequency as a probability that the sensor gathered and reported information on the social agents within their collection radius on any given simulation tick;
- The sensor's reporting accuracy as a probability that the information gathered by the technical sensor was accurate with respect to the location and group membership of observed social agents.

In addition to these three variables, a fourth variable was used to determine where the technical sensor would be placed on the landscape. This variable determined the location of the technical sensor using a similar method employed for determining the identity of human agents. In this case, a random draw between [0 1] was employed and then scaled to the total number of cells or patches on the landscape. Afterward, the resulting number was used to select a specific patch identification number, or "p.id," where the sensor would be located. For example, a random draw of 0.3402862 on an interval of [0 1], in a 50 X 50 landscape with 2,500 patches, would place a technical sensor at the patch with "p.id" 850.

Because technical sensors collect information differently than surveys or human agents it is important to understand how they aggregate information and support

inferences about the population of social agents. In this process, the technical sensor queries all of the agents within its radius, who report back their individual ids, group membership, and location with the specified likelihood of being truthful or deceptive. Based on this information, the sensor creates a comprehensive picture of the landscape out of the individual responses of the queried agents. In this process, each agent's provided group and location is checked with other responses. If two agents responded with different group identification, and provided locations that are neighboring, the sensor infers that they reside in a mixed neighborhood. This process differs from other forms of collection because it reconstructs the percentage of agents in mixed neighborhoods based on fusing identity and location data, rather than relying on direct reports of each agent's "my.mixed?" variable as given to surveys and human agents.

Collection Demonstration

The simulation of collection capabilities occurred in two stages. The first step validated the design of collection algorithms by positing the existence of perfect sensors, i.e. surveys, human agents, and technical sensors that made no errors in their reporting, covered the entire population of social agents, and reported information to analysts at every time step. With such capabilities, the information available to analysts using these idealized sensors should exactly match the true state of the simulation. The second stage relaxes the sensor capabilities, limiting their collection accuracy, range, and reporting frequency. These imperfections approximate real-world sensors and reveal how different ways of collecting intelligence within the segregation model reveal systematic errors or biases of different types based on collection capabilities.

Ideal Sensors

The results of ideal sensors collecting information on a social population are shown in the figures below. First, Figure 7-1 depicts the initial and final state of the simulated population as it segregated over time. Afterward, Figure 7-2 depicts the information from a perfect survey compared with that of the population's true state for the percentage of the population in mixed neighborhoods, the percentage of the population that was happy, and the average number of moves made by social agents in the population. Figure 7-3 provides the same information for an ideal human agent. Finally, Figure 7-4 provides a comparison of a perfect technical sensor's estimate of the percentage of the population in a mixed neighborhood with the true state of the population. Importantly, no information from technical sensors was available on the percentage of the population that was happy or the average number of moves made by each agent.

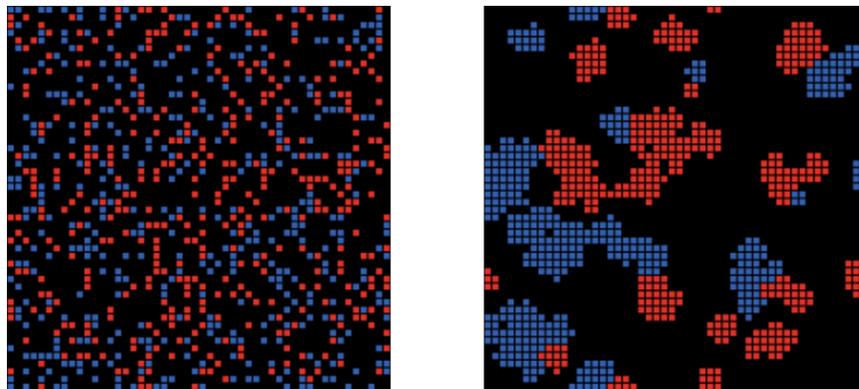


Figure 7-1: Initial and Final State of the Simulation. The landscapes above show the state of a population of 1500 agents on a 50 X 50 landscape. The image on the left is the agents at the start of the simulation and the image on the right shows the same population at the end, once each has moved into neighborhoods that satisfy their individual preferences.

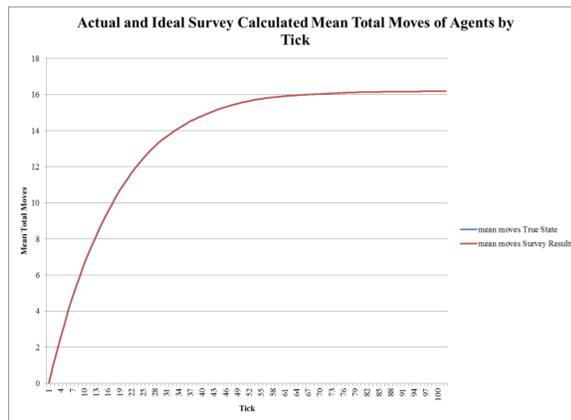
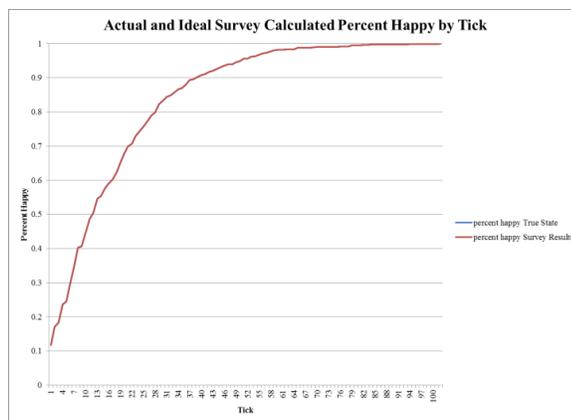
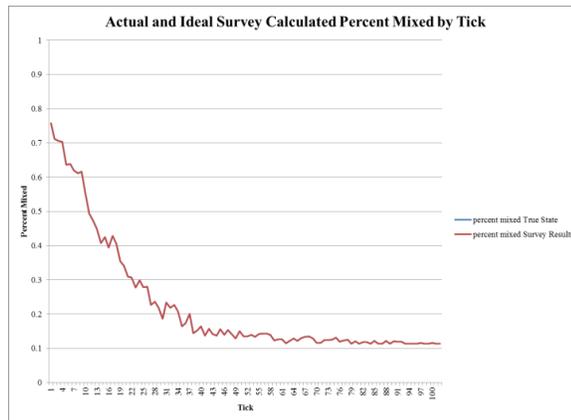


Figure 7-2: Comparisons between Ideal Surveys and the True State of an Agent Population. The images above compare ideal or perfect surveys with the true state of the agent population over the course of a simulation. In each case, the red line, data from the ideal survey, is indistinguishable from the blue line, the true state of the population. The fact that no differences between the survey data and the true state of the population can be found means that the methods by which survey information is collected are valid and correctly gathers and interprets agents' information via survey responses.

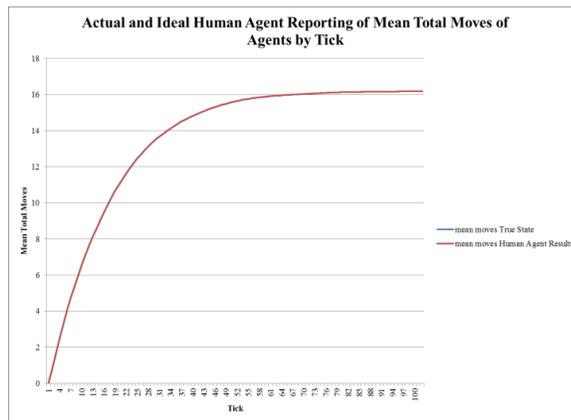
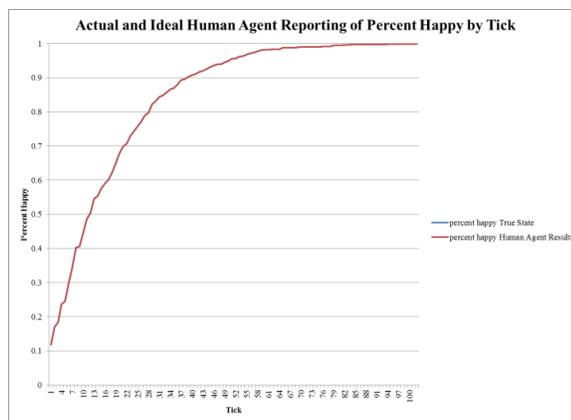
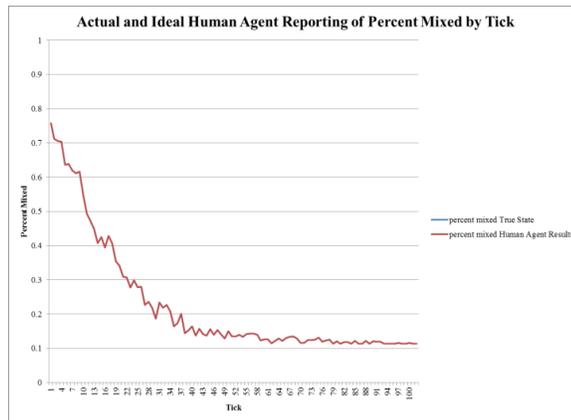


Figure 7-3: Comparisons between Ideal Human Agents and the True State of an Agent Population. The images above compare ideal or perfect human agents with the true state of the agent population over the course of a simulation. In each case, the red line, data from the ideal human agent, is indistinguishable from the blue line, the true state of the population. The fact that no differences between the human agent data and the true state of the population can be found means that the methods by which survey information is collected are valid and correctly gathers and interprets agents' information via human sources.

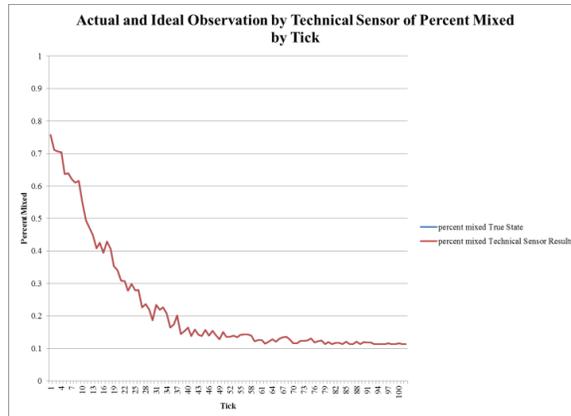


Figure 7-4: Comparison between a Perfect Technical Sensor and the Actual State of the Population. The image above compares an ideal technical sensor with the true state of the agent population over the course of a simulation with respect to the percentage of agents in mixed neighborhoods. Because the red line, data from the ideal technical sensor, is indistinguishable from the blue line, the true state of the population, the methods by which the technical sensor infers the percentage of the population in mixed neighborhoods are valid from a computational perspective.

In each of the graphs above, the sensor information shown in red is identical to the true state of the population depicted in blue (covered by the red lines and therefore unobserved). This is expected given the fact that the simulated sensors observed the entire population of agents accurately every tick. It validates the mechanism of information collection by sensors, including agent reporting and the fusion of their responses into a common picture. Because the information from ideal sensors matches the true state of the population, analysts can be confident that any differences that result from relaxing sensor performance are a result of changes in the performance parameters of the sensors and not artifacts of the sensor collection algorithms as implemented.

Flawed Sensors

The simulation of collection capabilities becomes more useful when sensors are known to possess flaws and imperfections, replicating the complexities of real-world information gathering from open and clandestine sources and methods. In order to demonstrate the ways in which alternative collection capabilities might be evaluated, a set of simple experiments was performed. In this simple set of experiments, the performance of 10 surveys, 10 human agents, and 10 technical sensors were compared with the true state of the agent population.

The specific parameters employed in each simulation are shown in Table 7-1 below. Because sensors operated independently of the population of social agents, each experiment evaluated individual sensors against the identical population, facilitating direct comparisons of their performance.

Table 7-1: Experimental Design for Sensor Simulation Demonstrations.

Parameter	Simple Experiments
red.agents	400
red.preferences	3
blue.agents	400
blue.preferences	3
total.surveys	10
survey.sample.size	0.03
survey.sample.frequency	0.5
survey.prob.longitudinal	0
total.technical.sensors	10
technical.sensor.range	0.1
technical.sensor.frequency	0.5
total.human.agents	10
human.sensor.range	0.1
human.sensor.frequency	0.5
survey.response.reliability	0.7
technical.sensor.accuracy	0.7
human.agent.reliability	0.7

Note. The “Simple Experiments” contain a set of experiments where ten approximately equivalent sensors of each type are compared with respect to their ability to collect three metrics regarding the identical model population and simulation path in the segregation model.

Simple Experiment Results

The results of the simulations are depicted in the Figures 7-5 and 7-6 below.

These graphs depict the metrics of the population’s percentage mixed and the percentage happy as gathered by the sensors and the true state of the population. Importantly, only the percentage mixed is estimated by all three sensor types, because happiness is not accessible to technical sensors.

The results show distinctive collection profiles for each type of sensor, while their relative capabilities regarding accuracy, sample size, and frequency are approximately equal. The graphs that follow show important structural properties, such as the tendency for collection to overestimate some measures while underestimating others, as well as sensor-specific structures of collection errors. Because this section's objective is to demonstrate the feasibility of employing ABM for simulating intelligence collection, the demonstration's data is examined visually and not subjected to more rigorous quantitative analysis because such depth is excessive given the simplicity of model and collection systems as represented. Although ABMs provide the potential for virtually unlimited quantitative investigation of simulation data when such precision is warranted, what matters for intelligence tradecraft and relations between analysts and collectors is that they can still gain useful insights from simple, cursory assessments of simulation results that establish a middle ground between data-free assessments of collection strategies and capabilities, and highly specialized and time-consuming quantitative analysis that requires specialized methodological or analytic support to examine and interpret. Given that the current relationship between analysts and collectors has been described by intelligence professionals as *ad hoc*, it would be unlikely that this process would warrant the attention and inclusion of large numbers of quantitative experts and specialists needed to mine volumes of simulation data. Therefore, this middle ground, where simulations can be developed and assessed in a cursory fashion, constitutes a significant contribution to tradecraft because it enables broad participation and uses fewer scarce, specialized

resources that can always be brought to bear if needed based on the specifics of the intelligence problem and simulation models employed.

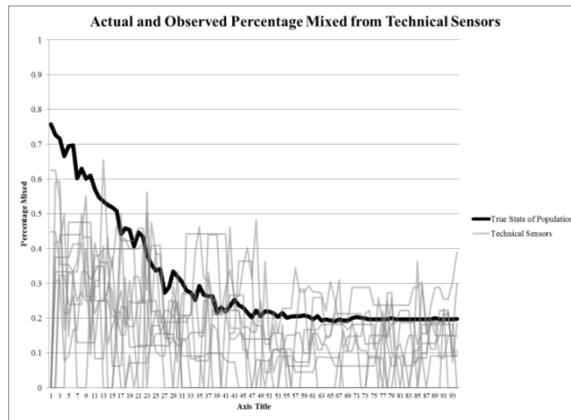
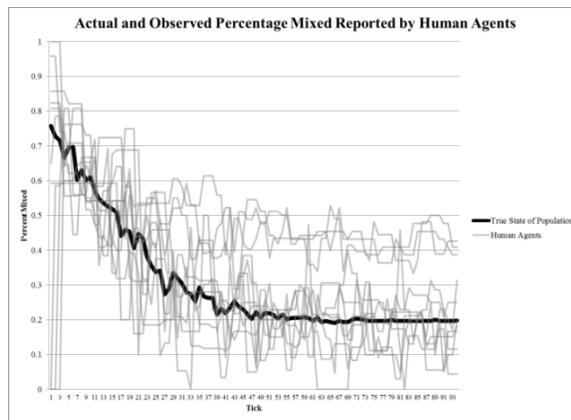
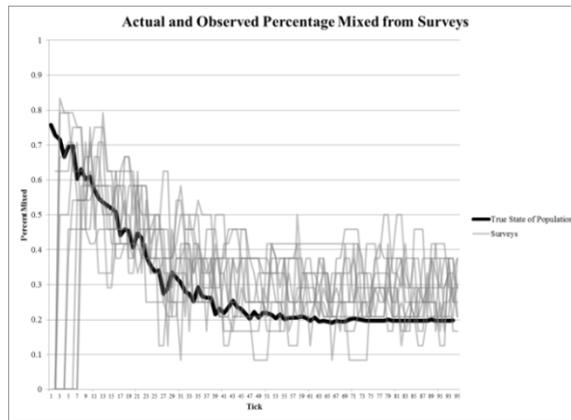


Figure 7-5: The Actual and Observed Percentage of Agents in Mixed Neighborhoods. The top image shows the percentage mixed as estimated for 10 surveys in grey compared with the true state of the population in black. The center image shows the same data for human sensors in grey compared with the true state of the population. The bottom image shows the percentage of agents in mixed neighborhoods as estimated by technical sensors in grey and the true state of the population in black.

Figure 7-5, preceding, shows the estimated percentage of agents residing in mixed neighborhoods for each of the three sensor types that was simulated. The percentage of agents residing in mixed neighborhoods is the most important metric to compare alternative sensor types for two reasons. First, it serves as the most important indicator of social mixing and the development of pluralistic communities free from strong ethnic, sectarian, or other biases, and is therefore the centerpiece of the segregation model's theory and social narrative. Second, it is the one variable that all three sensor types estimate, allowing for each to be compared. As noted earlier, technical sensors cannot provide information on the happiness of individual agents or their movement history.⁴³³

By examining the percentage mixed, the first difference is apparent when comparing the survey results with the human agents and technical sensors. Because survey responses are a random sample and not spatially related, the percentage mixed reported in surveys tends to overestimate the true state of the population. This is strictly a statistical phenomenon, resulting from the fact that if agents provide a random, error-prone answer to a true or false question, such as whether they reside in a mixed neighborhood, a random response can only produce one error-inducing conclusion. If the majority of agents reside in segregated communities, and therefore correctly answer false to the question of residing in a mixed neighborhood, then a random answer will only distort the result if true is given. Because the majority of the agents in a random sample reside in segregated neighborhoods, deceptive answers are more likely to respond true

⁴³³ To say that technical sensors cannot provide estimates of these measures is meant to note that this information cannot be inferred from the way technical sensors have been implemented in the model. Different implementations of technical collection processes and sensor capabilities may make these variables reachable, further extending the model and its sophistication with respect to its ability to simulate collection capabilities.

and therefore the result is overestimating the percentage of agents in mixed neighborhoods.

The characteristics of human agents and technical sensors are quite different from surveys. This is because their samples are not random but based on the location of agents in space. By gathering information on populations based on their spatial locations, rather than at random, response values will be contingent on the extent of spatial autocorrelation in neighborhood structure within the population, which is naturally high as more and more agents reside in segregated neighborhoods, thus having high likelihoods that agents will have neighbors that look like themselves. Simple visual inspection of Figure 7-5 shows that surveys are highly unlikely to underestimate the percentage of agents in mixed neighborhoods, while human agents and technical sensors produce errors of different types: Human agents displayed a capability of overestimation and underestimation, while the technical sensors showed a higher propensity to underestimate the percentage of agents in mixed neighborhoods (although not exclusively so).

In the simulation, human agents and technical sensors share an important commonality with respect to their spatial collection process. However, they also differ with respect to their placement in space on the duration of the simulation. Because human agents are mobile, and in fact follow a movement pattern that is identical to the rest of the population's agents, they will always move toward areas where they will interact with other agents, ensuring that samples are never empty (even if they only report on themselves in the worst of conditions). Alternatively, technical sensors are stationary and tied to the landscape. Therefore, it is possible that agents may move into and out of

their view, and they may even end up in locations where no agents are within their collection radius. Thus, these spatial collection capabilities, human agents and technical sensors, show different profiles in the calculation of percentage mixed variables.

The estimate of the percentage of agents in mixed neighborhoods by human agents shows an important sensitivity that reveals the significance of careful targeting in collection programs. Three grey lines that track the percentage of agents in mixed neighborhoods for human agents are noticeably and consistently higher than the other sensors and the solid black line that reflects the true state of the population. These lines do not reflect a simple error in the accuracy of the human reporting, but rather reveal how the selection of a particular human source can lead to a distorted understanding of the environment. The consistency of these errors suggests that the human sources reporting on their environments have themselves experienced a path through the simulation that is far from the mean with respect to their exposure to mixed neighborhoods.

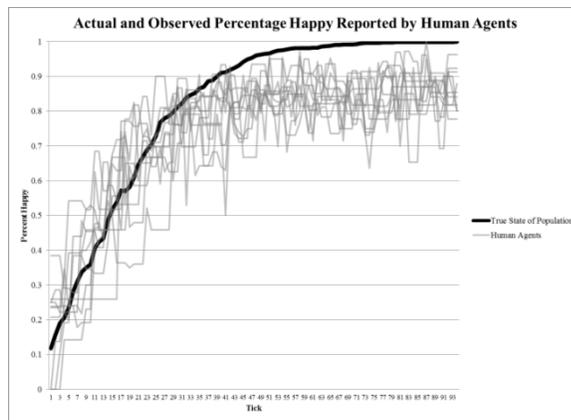
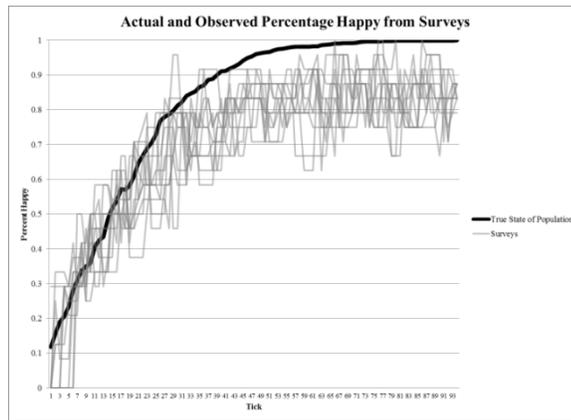


Figure 7-6: Actual and Observed Percentage of Agents Happy with their Current Neighborhood for the Surveys and Human Agents. The top image shows the results estimating the percentage of agents in mixed neighborhoods from 10 independently performed surveys in grey compared with the true state of the population shown in black. The bottom image shows the result of 10 independent human agents reporting on the percentage of agents in mixed neighborhoods in grey and the true state of the population in black.

The examination of the percentage of happy agents within the population shows additional properties regarding intelligence collection. First, it is important to note that only two of the three types of sensors have gathered information for estimating this metric: Technical sensors cannot access the internal states of agents and therefore cannot determine happiness. Second, surveys maintained a fixed sample size that was

independent of space while human agents collected information spatially based on the agents within the collection radius of the source, resulting in variable sample sizes from tick to tick. Because the collection algorithm, i.e. the ways in which surveys and human agents gathered information about the social agents in the model, was the same, all differences between their relative performance can be attributed to their treatment of space and the stability of their sample sizes.

Surveys and human agents both showed similar estimation profiles: Each overestimated and underestimated the percentage of the population that was happy until approximately tick number 40; afterward both systematically underestimated the percentage of happy agents. The crucial difference between the two sensor types was the surveys showed less variance in their estimates and therefore smaller errors. The smaller errors on the part of surveys can be attributed to the stability of their sample size. By comparison, because the number of agents within observation range of the human sources varies considerably, the volatility of the results is larger and therefore larger errors can be found. Moreover, the spatial character of human source collection can also produce higher errors because the results may be driven by the local experiences of communities on the landscape that may not reflect the state of the population as a whole; therefore, the global reach of surveys provides another advantage for computing the measures of the percentage of happy agents in the population.

At the 40th tick, an interesting pattern emerges that is independent of sensor type. Surveys and human agents both started to consistently underestimate the percentage of happy agents. This is due to the statistical factor noted early—that responses are binary

and the majority of agents had achieved happiness by the 40th tick. Because most agents were happy, the majority of deceptive responses would therefore report being unhappy, thus leading to an underestimation of the state of the population. Thus the underestimation of the population's happiness is a structural property that occurs independently of sensor type.

These experiments demonstrate the feasibility of uncovering the properties of different collection capabilities when employed against particular intelligence targets. These targets may be large social systems as performed here, or richly detailed models of specific organizations, facilities, etc. By simulating alternative collection strategies, analysts and collectors can work together to home in and match collection capabilities with analysts' information needs, while uncovering how biases might exist in data as a result of the structure of the targeted system.

Collection Portfolio Development

The evaluation of individual collection capabilities constitutes an important contribution of ABM to analytic tradecraft by facilitating and formalizing relations between analysts and collectors. However, using models of social systems to explore the behavior of specified collection systems only hints at the ways in which computation can aid analytic tradecraft. The use of genetic algorithms and other forms of evolutionary computation or machine learning can be employed to discover how to combine individual collection capabilities in order to optimize their use against specific targets, operate in specific scenarios, or identify combinations that may be robust across multiple cases wherever the adversary's actions or the behavior of populations may be unknown. As a

result, computation can penetrate the creative aspects of collection and analysis by searching for, and recommending, portfolios of capabilities whose specific configuration might not occur to collectors or analysts.

In order to demonstrate the potential opportunities afforded by ABM, the segregation model was modified again by adding two distinct layers to the simulation system. The first change was the use of the “sensor.portfolio” agent, introduced earlier, that owned a portfolio of individual sensors working together to develop a common, fused picture of the population. These sensor portfolios were constructed through the use of “tags,” i.e. strings of binary numbers, 0s and 1s, which characterized the type, capability, location, accuracy, and expense of the portfolio’s component sensors. Because all of the parameters that characterized sensors individually could be specified on the interval [0 1], implementing a generic tag structure that could accommodate all sensor types in a portfolio was simple from a software perspective. A simplified representation of the tag structure is shown below in Figure 7-7.

Tag representation of a single sensor in a portfolio. Each sensor is characterized by 41 tag positions that determine the sensor's parameters, such as sample size or collection range, location, accuracy, etc.

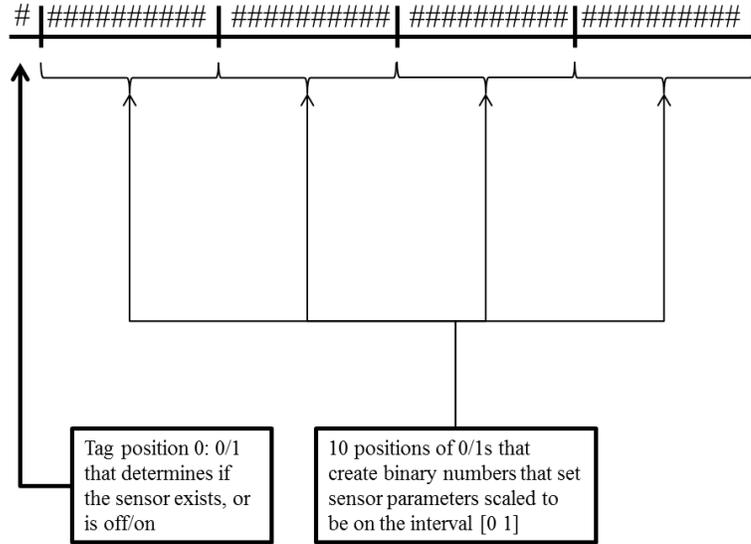


Figure 7-7: A Depiction of a Single Sensor in Tag Form. Each sensor in a portfolio is assigned 41 binary numbers that characterize its parameters. A sensor portfolio will contain several tags, one active for each of its sensors, i.e. with the first position set to 1. Sensors that are not active are set to 0 in their first position. The length of each portfolio's tag is based on the maximum number of sensors that portfolios are allowed to contain multiplied by three, the number of sensor types. As long as no more than the maximum number of sensors has their first value set to 1, the configuration is allowed and the portfolio will seek to fuse the information from each sensor into a comprehensive view of the social system. If a portfolio ever has more than the allowable number of sensors set to active, a random selection of sensors is set to 0, bringing the number of active sensors in the portfolio within the range of acceptable numbers again.

The second addition to the model was a simulation wrapper or experimental harness that repeated identical runs, ensuring that all social agents moved in an identical fashion every run. This process allowed for different collection portfolios to be compared against the true state of the social population and divided based on their relative performance or "fitness." Those with high relative fitness, i.e. at or above the

median of all portfolios, were retained in the population while those below the median fitness were discarded and replaced by new portfolios. These new portfolios were created by mixing the features of high-fitness portfolios based on the biological principles of genetic crossover and mutation. Specifically, the tags of high-fitness portfolios were combined to create new portfolios, and then modified at random by switching 1s to 0s and 0s to 1s with a specified probability, simulating sexual reproduction and genetic mutation in biology.⁴³⁴

Together, the development of sensor portfolios and their evaluation through the use of a genetic algorithm can aid analysts and collectors in developing intelligence collection strategies that are tailored toward analysts' particular needs. By selectively adjusting how fitness is defined, genetic algorithms can be employed to identify tradeoffs in the design of collection strategies and capabilities, allowing analysts and collectors to tailor the pursuit of information in order to address their intelligence needs—whether they are attempting to isolate specific sources of information, cross-check intelligence from multiple sources in order to detect the presence of adversary denial and deception activities, develop strategies for hedging against the weaknesses of individual collection systems, etc.

Genetic Algorithm Demonstration

A simple example of a genetic algorithm for examining collection portfolios is described below. Given the abstract level of the social system characterized in the

⁴³⁴ This technique is discussed in greater and more sophisticated detail in Melanie Mitchell, *An Introduction to Genetic Algorithms* (Cambridge, MA: MIT Press, 1998); and Ken De Jong, *Evolutionary Computation: A Unified Approach* (Cambridge, MA: MIT Press, 2002).

segregation model and the simplicity of the sensors described by the model, the results of this demonstration should not be interpreted as applying to specific collection systems or targets. Instead, it demonstrates the viability of employing evolutionary computation in tandem with ABM within the context of analyst–collector relations, and provides some insights into general features of collection as it relates to the analysis of complex social systems.

Implementing a genetic algorithm for exploring intelligence collection required the additional modifications to the segregation model noted above. These modifications included the development of new agent classes for tracking the performance of sensor portfolios, and an experimental harness for managing populations of these portfolios while repeating the same historical path of agent movement and settlement in the social population. Because the central concept behind the development and use of genetic algorithms is allowing populations of “solutions”—in this case sensor portfolios—to discover new ways to solve problems, a basic scenario established by the running of the segregation model needed to be repeated multiple times in order to allow for the evolutionary processes of crossover and mutation to evolve new, higher fitness portfolios over multiple generations. This required maintaining a technical division between the random number generators that governed the movement of social agents, and those that belonged to each sensor portfolio.

As noted at the start of this section, ideal sensors could collect against a target population in an error-free way. However, collection systems of this kind do not exist in the real world due to technical limitations of collection systems, strategic behavior on the

part of targets, and, even if they were possible, the costs of engineering and operating systems to such tolerances would be prohibitive. In order to prevent the genetic algorithm from “discovering” these perfect sensors by moving into a corner solution where a sensor portfolio’s tag consisted of all 1s for its sensor capabilities, the possible parameters of collection sensors were “throttled,” capping certain collection capabilities such as the sample size of surveys, the reporting range of human agents, or accuracy of technical sensors to a particular threshold. Each of these parameters was reduced by a specified percentage in order to prevent the emergence of singular, perfect sensors. Thus, a throttle factor of 0.8 for a survey’s sample size would multiply the value of this parameter created by its tag by 0.8, effectively throttling the sample size by 20 percent.

Another feature of the genetic algorithm was the development of a sensor fusion capability for combining the results of multiple sensors into a common view of the population. Given that each sensor possessed distinct properties, including the likelihood that it might report incorrect information, and different collection properties and processes, combining intelligence gathered from each individual sensor into a single perspective on the society was non-trivial. For example, a portfolio might have identified “agent 32” from multiple sensors, where one placed it at “patch 34,” another sensor located it at “patch 37,” and still a third located the agent at “patch 19.” Whenever confronted with multiple, conflicting reports, the portfolio resolved conflicting information by choosing a reported value at random. This simple scheme biased the development of a common perspective toward selecting those values that occurred with the greatest frequency for each observed data element. More sophisticated treatments of

this problem might have included correlating observations across time, developing histories of sensor performance to weigh individual reports, and cross-checking reported values with sensor properties, e.g. discarding the reports from a spatial sensor if the observed location for an agent is outside of its collection radius.

This sensor fusion capability served as the basis for the third feature of the genetic algorithm: the design of the fitness function. There were many ways to determine fitness, and the features of solutions to problems can vary greatly depending upon how it is specified. In the case of the demonstration, fitness was defined by adding two metrics together: the accuracy of the portfolio at the microlevel and macrolevel of analysis.

The microlevel fitness of the sensor was determined by the number of features the portfolio got correct about each individual agent it observed. The microlevel features that factored into the sensor portfolio's fitness function included:

- Whether the portfolio associated the correct individual id to the agent, e.g. whether the individual identified as “agent 32” was actually “agent 32”;
- Whether the sensor portfolio located the agent in its actual location or incorrectly assigned it to a different location;
- Whether the sensor portfolio assigned the agent to its correct group;
- Whether the sensor portfolio reported the correct number of total moves made by the agent;
- Whether the sensor portfolio correctly reported the last time the agent had moved locations;
- Whether the sensor portfolio correctly reported whether the agent was happy;

- Whether the sensor portfolio correctly reported whether the agent was in a mixed or single group neighborhood.

The macrolevel fitness of the sensor portfolio was determined by comparing the portfolios' estimates of several distributional features of the population with the actual values of the population. These measures included:

- The estimated percentage of the population belonging to the red group or the blue group;
- The estimated percentage of the population that is happy;
- The estimated percentage of the population that is in a mixed neighborhood;
- The estimated mean, median, min, max, and standard deviation of the total moves within the agent population;
- The estimated mean, median, min, max, and standard deviation of the voluntary moves within the agent population;
- The estimated mean, median, min, max, and standard deviation of the forced moves within the agent population;
- The estimated mean, median, min, max, and standard deviation of the last tick moved by agents within the population;
- The estimated mean, median, min, max, and standard deviation of the number of neighbors of agents within the population.

Evaluating sensor portfolios at two levels of analysis was necessary in order to balance alternative collection strengths. For example, a sensor may have correctly gathered information on individuals, but if that sample was biased or confined to a small

geographic region of the system, it may have produced highly misleading estimates about the larger society. Likewise, a portfolio may have been highly inaccurate regarding individual agents, but nevertheless provided accurate information about the larger population, e.g. the case in which equal numbers of red agents are reported as blue, and blue agents are reported as red, would yield low microlevel accuracy, but nevertheless provide an accurate macrolevel, systemic assessment of the population. While it is more likely that microlevel accuracy will lead to macrolevel accuracy, this should not be assumed given the presence of highly skewed distributions of variables within the population and spatially disconnected and segregated subgroups that may be missed by small random samples of the population or geographically confined collection.

Figure 7-8 below provides an indication of how a genetic algorithm was used to develop new collection portfolios, and identify what features are best suited for collecting information about the population of social agents within the basic segregation model. In this example, a population of 30 sensor portfolios was evolved for 30 generations, with each portfolio possessing a maximum of 4 sensors fused together in order to identify the best sensor type, or combination of sensor types. While a complete analysis would require many more generations and a larger population in order to ensure a more complete search of possible portfolio configurations, this small population nevertheless generated evolutionary dynamics that were sufficient for demonstrating their potential to analysts and collectors alike.

Over the course of 30 generations, the population of sensor portfolios evolved from having an initial average fitness of 0.345687784, to an average fitness of

0.688088794. More importantly, the peak maximum fitness within the population of portfolios was initially 0.550982469, and grew to 0.693678605 after 30 generations. While the numerical values are not significant, the general trend they reveal is. The computational machinery responsible for managing the simulation learned to improve the portfolios of sensors within the model, and worked toward identifying solutions specifically tailored to gather information in the specified scenario. Figure 7-8 below shows several metrics of the genetic algorithm's operation for the initial generation 0, the midpoint of generation 15, and the final generation 30. Each panel shows how the population of sensor portfolios changed over time, demonstrating how collectors and analysts might capitalize on computational machinery in their search for strategies that match collection capability with the informational needs of analysts.

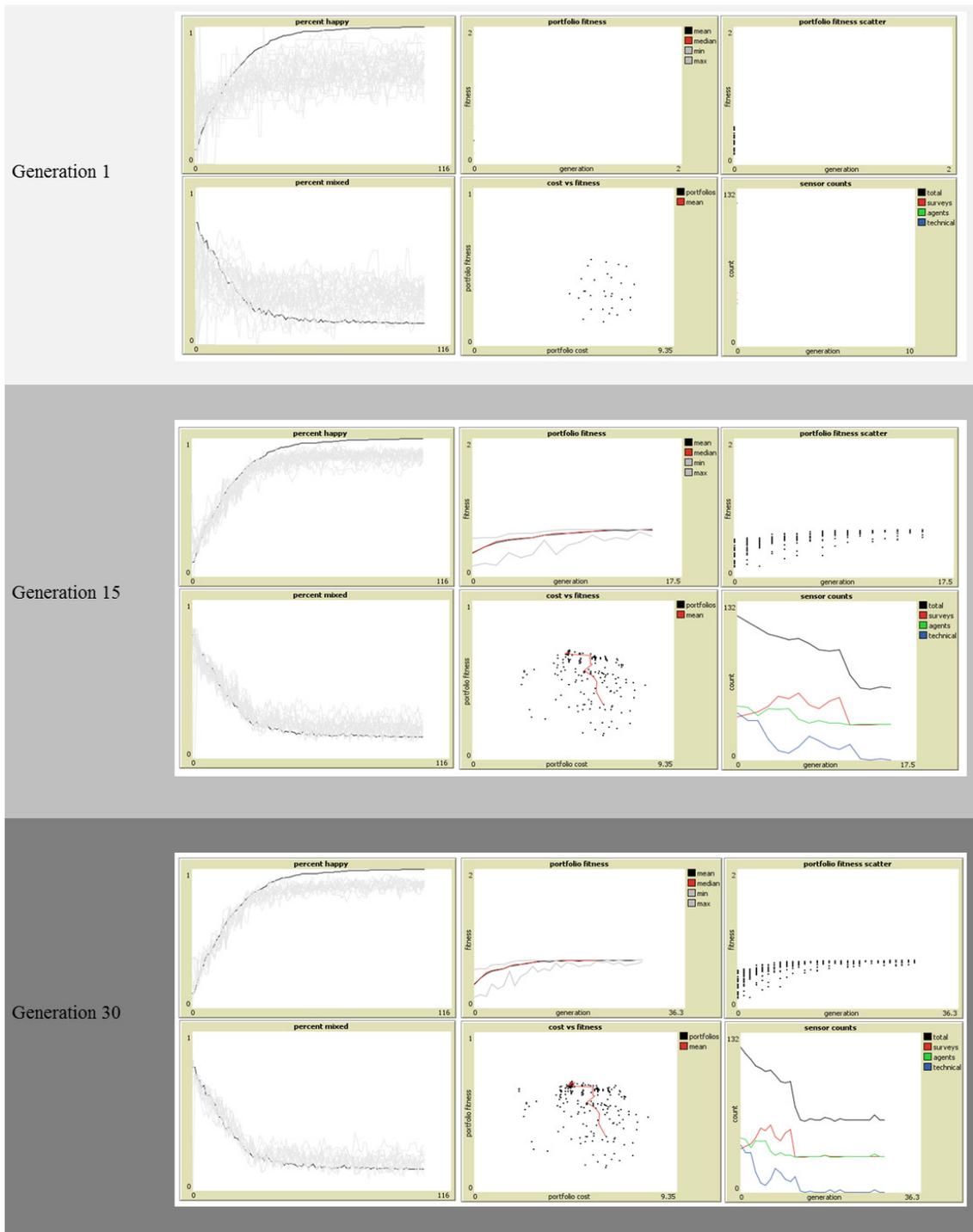


Figure 7-8: Panels Depicting the Status of a Genetic Algorithm for Developing Sensor Portfolios. Each panel depicts three different generations of the genetic algorithm, revealing the journey from an initial sample of randomly created portfolios into a closely knit collection of high-performing options specifically tailored to collect information about the observed social population of the segregation model.

The panels in Figure 7-8 reveal the history of the genetic algorithm and the ways in which it improved the population of sensor portfolios from one generation to the next. The most obvious improvements are shown in the two graphs on the upper and lower left of each panel, shown in greater detail in Figure 7-9 below.

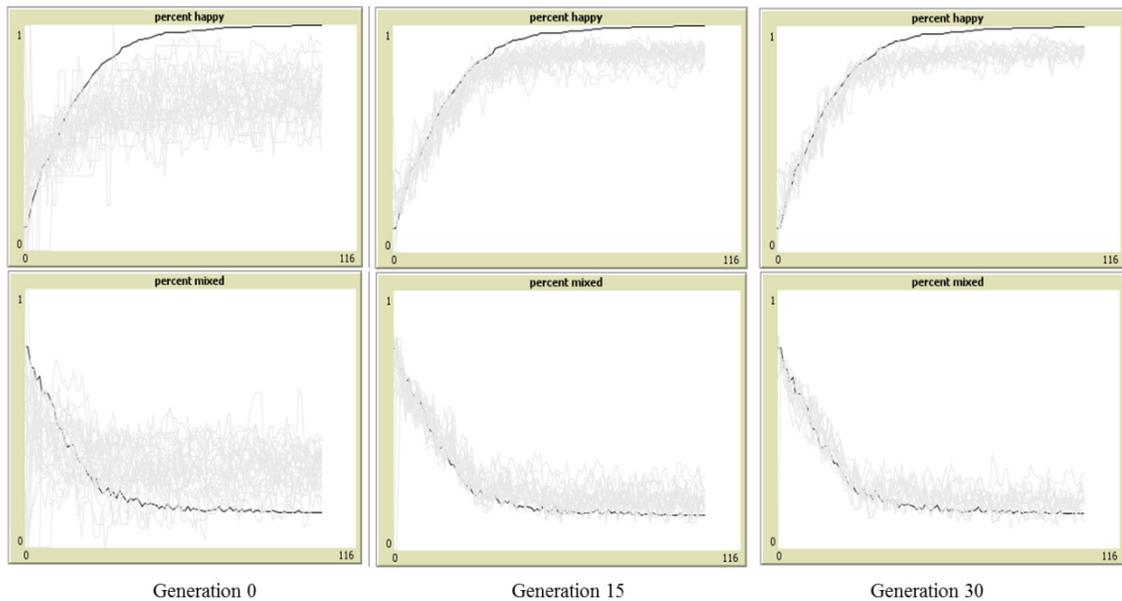


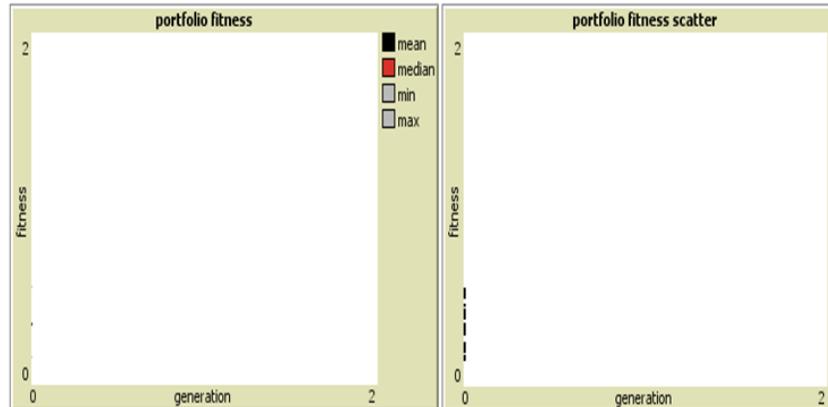
Figure 7-9: The Percentage Happy and Percentage Mixed Metrics as Estimated by Sensor Portfolios for Generations 1, 15, and 30. The light grey lines depict the reported values of each sensor portfolio. The solid black line shows the true state of the agent population being observed by the sensors.

The graphs in Figure 7-9 show the true percentage of agents that are happy with their local neighborhoods for each tick of the simulation and the percentage that were in mixed neighborhoods—the same metrics discussed earlier in this section when examining sensors individually. The single black lines depict the true state of the agent population,

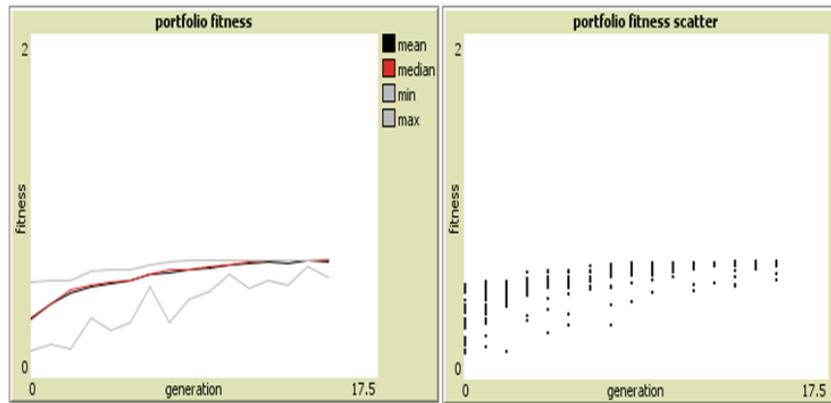
while the grey lines show the estimates of these variables from each of the individual 30 sensor portfolios in the population under examination for that given generation. At generation 1, the population of randomly generated portfolios shows significant errors when compared with the actual values, as shown by the distance and spread of the gray lines. With each passing generation, however, these lines show less variability and better approximate the true value shown by the solid black lines for both the percentage happy and mixed. Importantly, even though the sensor portfolios clearly improve over time, they continue to underestimate agent happiness and overestimate the percentage of the population in mixed communities for the reasons noted earlier.

Additional information about the population of sensor portfolios is provided in the four additional graphs in each panel. The top center and right graphs provide two perspectives on the calculated fitness of the population of sensor portfolios, and are shown in greater detail in Figure 7-10.

Generation 1



Generation 15



Generation 30

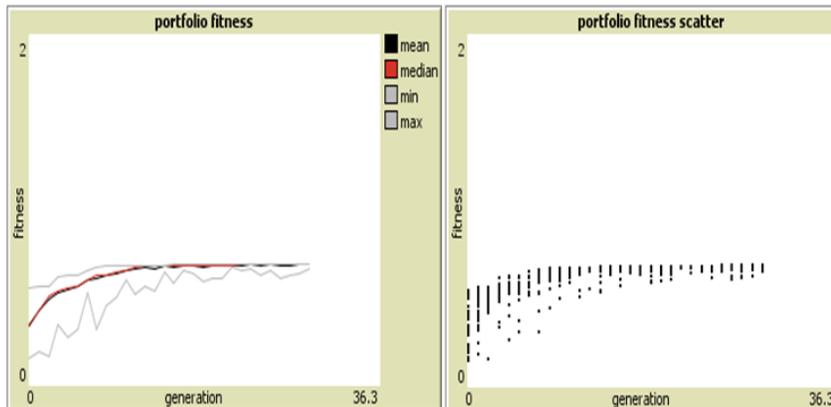


Figure 7-10: Progression of the Population of Sensor Portfolios. The plots above show the major statistical properties of the portfolios' fitness values for their initial, fifteenth, and thirtieth generations on the left panels. The right panels show the distribution of fitness values for each generation, providing greater detail into the diversity of the portfolios being evaluated.

These images in Figure 7-10 depict how the population of sensor portfolios improved their fitness over time. The panes on the left show the min, max, mean, and median fitness of the sensor portfolio population for generations 1, 15, and 30. These lines show two significant trends: There was an initial, rapid jump in the fitness of the population that occurred as ineffective portfolios in the initial population were quickly identified and replaced by the offspring of higher fitness portfolios. Afterward, a prolonged and slow ascent in fitness occurred, corresponding with a decrease in the relative diversity in the population of sensor portfolios, shown by the narrowing gap between the min, max, mean, and median fitness values. This story is confirmed by the graphs on the right of Figure 7-10, which provides the same information in the form of a scatter plot, where the vertical height of each point depicts the fitness of a single portfolio while the horizontal position corresponds to the generation. Again, the early generations of the genetic algorithm showed single points that spanned a wide vertical range, which rapidly diminished in terms of vertical range and rose with respect to vertical position with each passing generation.

The journey of the portfolios from positions of low fitness to high fitness with respect to their ability to collect information on the target population was also depicted in relative terms—specifically, as a comparison between the relative fitness and cost of each portfolio—thus creating a richer understanding of how improvements in the collection capabilities of each generation of portfolio occurred. This journey is shown in Figure 7-11 below.

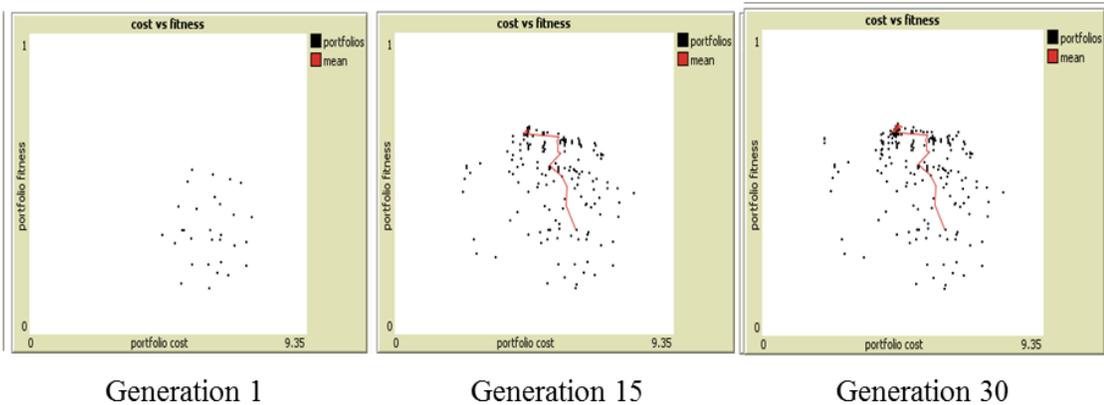


Figure 7-11: Scatter Plots Depicting the Portfolio Cost and Fitness for Multiple Generations. Each point depicts the cost of the portfolio on the x-axis and the fitness of the portfolio on the y-axis. The red line tracks the median cost and fitness of each generation in order to observe how the population of sensor portfolios are changing over time.

The points in black shown in each of the graphs in Figure 7-11 constitute the location of each sensor portfolio in a space defined by their relative costs and fitness. The red line depicts the mean cost and fitness of the portfolios in each generation, allowing for a clearer visualization of the path taken by the genetic algorithm over each generation. The evolution of the sensor portfolios showed a clear trajectory, where the red line started moving up and left, indicating improvements in overall sensor fitness and a reduction in cost over the first several generations, most clearly shown by the differences between Generation 1 and Generation 15. The change from Generation 15 to Generation 30, however, showed a different dynamic, resulting from the increased homogeneity between population members. Over this period, new portfolios were highly similar to existing ones, and thus few changes in cost and fitness resulted, leaving the mean values of the population unaffected.

The movement trajectory also reveals the search path of the algorithm, where the tendency is for the mean to move up, indicating improved fitness, or left, indicating a reduction in the cost of the sensor portfolios. In the early generations, the movement up in fitness was dramatic, while later generations primarily moved to the left, thus maintaining collection accuracy while reducing the overall costs of the portfolios for reasons that became evident in Figure 7-12, which tracked the overall number of sensors in the population of portfolios.

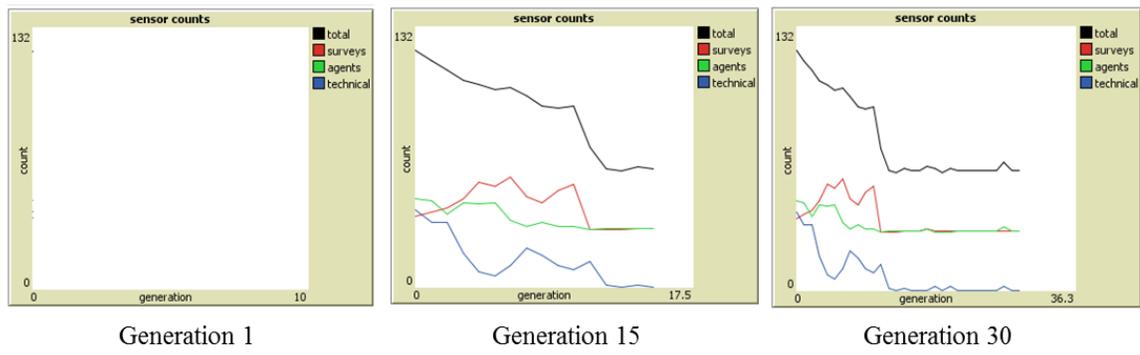


Figure 7-12: Sensor Counts by Type for the Population of Portfolios by Generation. The black line shows the total number of sensors in the population of portfolios. The red line shows the number of surveys, the green shows the number of human agents and the blue lines shows the number of technical sensors in the population.

Figure 7-12 shows the total number of sensors held by the population of sensor portfolios in black, and the surveys, human agents, and technical sensors in red, green, and blue respectively. While these graphs do not reveal the specific properties of the best sensor portfolio, they do reflect which sensor types are the most used in portfolios. In the case of the sample population, technical sensors were quickly eliminated and only

reappeared occasionally based on the mutation of new tags. Their elimination from the population accounts for the dramatic reduction in cost shown in Figure 7-11.

Importantly, even though portfolios were allowed to possess a total four sensors each, the actual number of sensors held by the population of 30 portfolios was far below the maximum and initial value of 120. Instead, the population settled into a state where the total sensors in the population remained around 60, averaging two sensors per portfolio, and the total types of sensors within the population were equally divided between surveys and human agents. This was due to the fitness function considering dimensions of information such as agent movement history and happiness that were not accessible to technical sensors, thereby reducing their fitness relative to the other two sensor types.

At the conclusion of the 30 generations, the highest performing sensor portfolio possessed a cost of 4.484848484848484 (calculated by the expense of the capabilities of its component sensors) and was composed of a single survey and a human agent. The survey sampled 70 agents out of a population of 800, with an accuracy of 78%, a reporting rate of 66% (a probability of 0.66 for collecting and reporting information each time step), and a 17% chance of tracking members of its sample longitudinally. The human agent reported on a radius of 5 units, on a 50-by-50 grid, with a reporting frequency of 75% and an accuracy of 77%. Together, these two sensors provided the most accurate information about the population at the macro- and microlevels when compared to all of the examined configurations.

The behavior of the genetic algorithm is consistent with its underlying biological metaphor where poorly performing portfolios were eliminated from the population and

replaced with the offspring of those with higher fitness. Thus, successful portfolios quickly dominated the population, simultaneously raising the overall level of fitness of the population while reducing its diversity at the same time. Additional improvements to the population then rested on the operation of crossover and mutation on portfolio tags, slowly uncovering new settings that outperformed previously known ones.

Summary

The demonstrations in this chapter argue for the use of ABM as a tool for coordinating relationships between intelligence analysts and collectors. As the first demonstration showed, ABM allows for the simulation of complex social systems and processes, while simultaneously affording opportunities to examine the performance of real or potential intelligence collection capabilities and strategies. Likewise, the second demonstration built on the notion of simulating intelligence collection by extending the use of computation to extend the exploration of collection possibilities. In doing so, the problems posed by satisficing in the search for collection options can be overcome by allowing the computational machinery to search large numbers of cases and combinations that could not be done manually.

Each of the capabilities demonstrated in this chapter are not unique to ABM. Operations research and other techniques have been employed in the past to simulate intelligence collection against particular targets, and genetic algorithms and other machine learning techniques have long served as a means for optimization in engineering and decision support. However, in each case, ABM provides an additional layer of capability that is difficult to achieve in its absence. More specifically, ABM allows for

the simultaneous specification and exploration of a complex social system, while simulating the collection of information about that system. This means that analysts and collectors can examine the interface between the behavior of the target and the means for understanding that behavior without holding one or the other constant and fixed. As intelligence targets become increasingly distributed, networked, adaptive, and often abstract—where threats are not found in the behavior or choices of single targets but in the ways in which actors interact with one another—the need for simultaneous examinations of collection opportunities and vulnerabilities, and alternative specifications of the intelligence target itself, is a necessity.

Chapter 8: Structure and Agency

Earlier chapters noted that intelligence analysts confront questions that require thinking about the relationship between structure and agency in order to determine how much the individual choices of foreign decision makers and consumers shape the international system. One of the benefits of ABM is its ability to generate macroscopic social phenomena from the actions and interactions of individuals, providing model users with the ability to identify and trace relationships between decision making, random shocks, and systemic outcomes. However, most of the experimental designs and analytic tools for exploiting ABM have not been tailored to address the fundamental needs of intelligence analysts and policymakers, leaving ABM underdeveloped and underexploited within the context of strategy and policy.

Modifications for Microscience Experimentation

The idea of a microscience for intelligence was discussed in Chapter 4. At its core was the need to examine systems from the perspective of individual choices and action in order to tease out the extent to which human agency can affect the international system or social systems more generally. While ABM provides a means for exploring questions of this kind, modelers have not committed to designed models to facilitate such explorations or develop the experimental infrastructure for doing so. Indeed, microscience is significant within the context of strategic theory, where actors seek to

identity opportunities and risks where small efforts or differences can have amplified consequences, as noted in the classical proverb of the king losing his kingdom due to a missing nail.

For want of a nail, the shoe was lost;

For want of a shoe, the horse was lost;

For want of a horse, the rider was lost;

For want of a rider; the battle was lost;

For want of a battle; the kingdom was lost!⁴³⁵

The third demonstration with the segregation model demonstrates how ABMs can be instrumented and experimented with in order to examine counterfactuals and the macroscopic consequences of individual choices and actions at the microlevel. In order to do this, several technical changes to the core model described in Chapter 5 were made, the most important of which altered the ways in which agents moved by giving each a unique random number stream. By developing an experimental design where each agent behaved identically, with the exception of a single agent that made different choices from run to run which required isolating each agent in software, and an ability to consistently access shared resources, such as patches on the landscape, without affecting the movements of other agents. While this introduced several computational inefficiencies, it provided important flexibility that was not available in the original model.

The following demonstration, then, discusses two alternative versions of the segregation model. The first is a technically modified version of the core model, but is

⁴³⁵ James Gleick, *Chaos* (New York, NY: Penguin Books, 1987), p. 23.

theoretically identical to the original version specified by Schelling. The second version, referred to as purposeful, altered the logic of agent movement in order to allow agents to adapt to one another's movements. This change was necessary in order to avoid overestimating the extent to which agents' choices impacted the system by accounting for the willful activities of others and their ability to adjust to changing circumstances. Together, each version demonstrates the feasibility and difficulty of achieving the objectives of microscience through the use of ABM.

Random Versus Purposeful Movement

The centerpiece of microscience is exploring how small changes in the behavior of agents can affect the entire structure of systems. In the segregation model, the behavior in question is concerned with when and where agents decide to move. In the baseline model, unhappy agents move to a randomly selected, empty patch on the landscape. If they attempt to move into a patch that is already occupied, the agent simply tries again and looks for another empty patch, repeatedly selecting patches at random until a vacant one is selected.

The default movement rule takes no interest in why an agent decides to move to one patch or another, and assumes that the agent is agnostic as to its location and only cares about its neighboring agents once settled. Because movement is entirely random, if an agent's first effort to find a new patch is blocked, any subsequent, randomly selected patch is assumed to be just as good from the perspective of the moving agent. Movement is not purposive and does not account for how agents might adapt to one another's

movements because there is no spatial component to the selection of target patches, effectively decoupling the decision to move from where to move to.

To address this question of strategic interaction and adaptation, a new movement rule was introduced that assumed agents' decisions regarding where to move contained some intrinsic motivation. While this still assumed that unhappy or coerced agents selected new locations at random, they adapted in a willful fashion if their initial choice was blocked, thus maintaining a sense of spatial awareness in their movements. As a result, rather than continue to seek out random unoccupied patches if their initial choice was occupied, purposeful movement instructed agents to move to the nearest empty patch from their initial choice.

Implementing these new movement rules required redesigning the model at the technical level by requiring that each agent possess a unique random number stream, as well as ways of selecting patches that were independent of the landscape's status. For example, the default movement rule capitalized on NetLogo's built-in geography and central random number stream, allowing agents to randomly select only patches that were unoccupied whenever they were required to move.

Isolating the individual choices and actions of agents required a new approach to movement and the use of shared resources, such as the patches on the landscape. Each agent required a means of selecting patches that only relied on its unique random number stream, rather than a shared one that would disrupt the decision making of other agents. This was achieved by giving each patch a unique identification number that was then accessed by drawing a random number on the interval $[0, X]$ where X is the total number

of patches on the landscape, e.g. a 50 X 50 landscape has a total of 2,500 patches with patch identification numbers running between 0 and 2,499. These changes effectively overrode NetLogo's built-in geometry and random number generator. Importantly, the need to ensure intra-run and inter-run consistency with random number draws for the purposes of movement required accessing patches via lists rather than NetLogo's default sets. As a result, identical random number draws by an agent would produce the identical patch id and ensure consistency throughout all simulations, which would not have occurred using NetLogo's default patchsets and agentsets.

Experimenting with Agency and Structure

In order to examine the relationship between structure and agency, the modified segregation model was examined under its normal movement rules and also under the new purposeful movement rule discussed above. This allowed for the identification of the consequences of agency and its relationship with structure by differentiating between those aspects of the system's structure that are immutable from those that are contingent on agents' choices. This experimental design established microlevel control over the agent population in ways that are rarely performed with ABM, and required isolating particular agents in the experimental process in order to examine how their choices affected the system as a whole, while holding the initial configuration of the model and the choices of all other actors constant.

These experiments were conducted in two groups. The first examined the modified version of the core segregation model, where movement remained completely random. Thus, the first group of experiments occurred with a technically altered model,

even though its theoretical representation was identical to the core model introduced in Chapter 5. The second group employed the purposeful behavior rule, where after randomly choosing a target location, moving agents searched for the closest unoccupied patch. This model was theoretically different than the original model because the logic of agent movement, particularly with regard to their intentions and considerations of space, was altered.

In each experiment, preliminary simulations were run one time to establish a baseline result, referred to run number 0. Then, the identical simulation was repeated 10,000 times with one exception: A single agent's random number generator was altered to provide it with a new random number stream in each simulation, holding all other agents constant. As a result, all differences between the baseline run and the subsequent 10,000 runs occurred because of the direct and indirect consequences of a single agent making different choices.

A total of fourteen experiments were conducted, where the choices of agents 0 through 6 were examined as described above under each of the two movement rules. The result was a total of 14,014 simulation runs that only scratched the surface of exploring the full dimensionality of artificial societies created by ABMs. The specific parameters explored in this demonstration are shown in Tables 8-1 and 8-2 below.

Table 8-1: Experimental Parameters for Normal Movement Experiments.

Normal Movement Experiments	1.0	1.1	1.2	1.3	1.4	1.5	1.6
Agents	1000	1000	1000	1000	1000	1000	1000
Preferences	3	3	3	3	3	3	3
Landscape	50 X 50						
Free Agent ID	0	1	2	3	4	5	6
Purposeful Movement	FALSE						
Setup Seed	100	100	100	100	100	100	100
Total Runs	10001	10001	10001	10001	10001	10001	10001
Baseline Run	0	0	0	0	0	0	0

The experiments in Table 8-1 show how the effects of individual agency were examined with the modified segregation model. In each case, the features and initial conditions of each simulation were identical with one exception: the identity of the “free agent” whose individual random number stream was altered in each simulation in order to allow it to make different choices from run to run. The same basic design was repeated using the purposeful movement rule as shown in Table 8-2 below.

Table 8-2: Experimental Parameters for Purposeful Movement Experiments.

Purposeful Movement Experiments	2.0	2.1	2.2	2.3	2.4	2.5	2.6
Agents	1000	1000	1000	1000	1000	1000	1000
Preferences	3	3	3	3	3	3	3
Landscape	50 X 50						
Free Agent ID	0	1	2	3	4	5	6
Purposeful Movement	TRUE						
Setup Seed	100	100	100	100	100	100	100
Total Runs	10001	10001	10001	10001	10001	10001	10001
Baseline Run	0	0	0	0	0	0	0

The final aspect of experimental design was the development of two analytic metrics for analyzing similarities and differences between simulations. These metrics quantified the differences between simulation runs in two distinct ways, each of which allowed for results to be compared across experiments. The first metric was based on creating a composite result of all the simulations performed in a given experiment. Afterward, each run in the experiment was compared against the composite to obtain a total difference measure in a fashion similar to a statistical standard deviation. The measure of distance from the average or composite result is referred to as “run.distance” and the process for computing it shown below.

$$Rd = \sqrt{\left(\sum_{Pr} \left(1 - \frac{Tr}{N}\right)^2\right) + \left(\sum_{Pb} \left(1 - \frac{Tb}{N}\right)^2\right) + \left(\sum_{Pe} \left(1 - \left(1 - \frac{Tb + Tr}{N}\right)\right)^2\right)}$$

Where:

- Rd = run.distance measure
- Tr = Total runs with a red agent on this patch
- Tb = Total runs with a blue agent on this patch
- N = Total runs in the experiment
- Pr = Patches with Red Agents
- Pb = Patches with Blue Agents
- Pe = Patches that are empty

Because the composite landscape established high-level probabilities that a given patch would be occupied by a red agent, a blue agent, or be empty, it implicitly assumed the existence of an average outcome or social result that was continuous. However, this assumption of continuity can be problematic for comparing discrete worlds, particularly if the average values are fundamentally unobtainable or cannot be found in the distribution of results. As a result, a secondary metric was developed that compared each run to a single baseline result, characterizing results as a percentage difference between the examined outcome and the baseline result with respect to the location of red agents, blue agents, and empty patches.

Together, these analytic metrics provide measurements of the differences between simulation runs, which only vary based on a single agent's choices, allowing analysts to

unpack those features of complex systems that are contingent on individual choices of agents and those that are immutable and embedded in the system's structure.

Analysis of Simulation Results

Exploring the relationship between agency and structure required assessing the segregation model in both of its modes, examining the simulation results based on the normal movement rule and the purposeful movement rule. The following discusses the results of the normal movement rules first, and then repeats the analysis with the purposeful movement rules in order to show how adaptive behavior by agents in the simulation limit the extent to which individual agents can influence the system as a whole.

Normal Movement

The normal model described above was examined under seven different scenarios where agents 0 through 6 were each allowed to make different decisions while the others were held constant. In all seven experiments, the initial conditions were identical and shown in Figure 8-1 below.

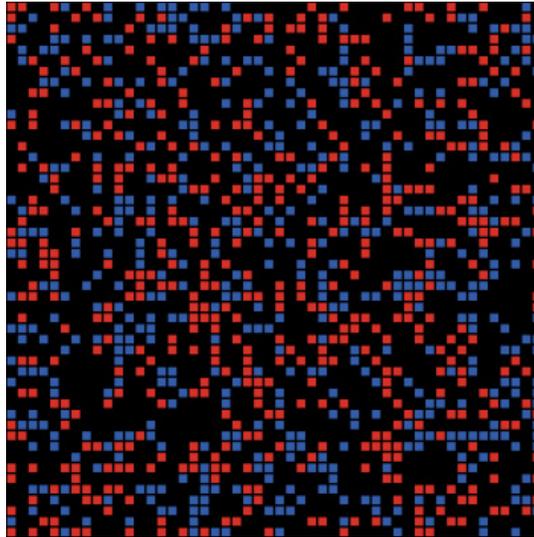


Figure 8-1: Initial Condition for the Segregation Model in All Experiments.

The initial position of free agents 0 through 6 are shown in Figure 8-2. Because these agents started in different positions, and moved in a different order, their relative influence over the system as a whole varied.⁴³⁶ The prospect that each agent may affect the system differently serves as a measure of their relative power and influence. This is significant because in all other respects, the agents in the segregation model are homogeneous with respect to social influence in that, in the absence of enforcers, none has direct control over others. By isolating agent power in terms of their relative influence over the system, it provides a more accurate assessment of the agents and the system itself that cannot be derived by studying their capabilities and intentions in isolation: As agents have one capability of moving to new patches and one intention, to move until they are happy with their local neighborhood.

⁴³⁶ In the simulations the activation order remained tied to Netlogo's central random number generator and was fixed for all experiments. As a result, the order in which agents moved was identical from run to run, placing agent activation into the category of a structural variable.

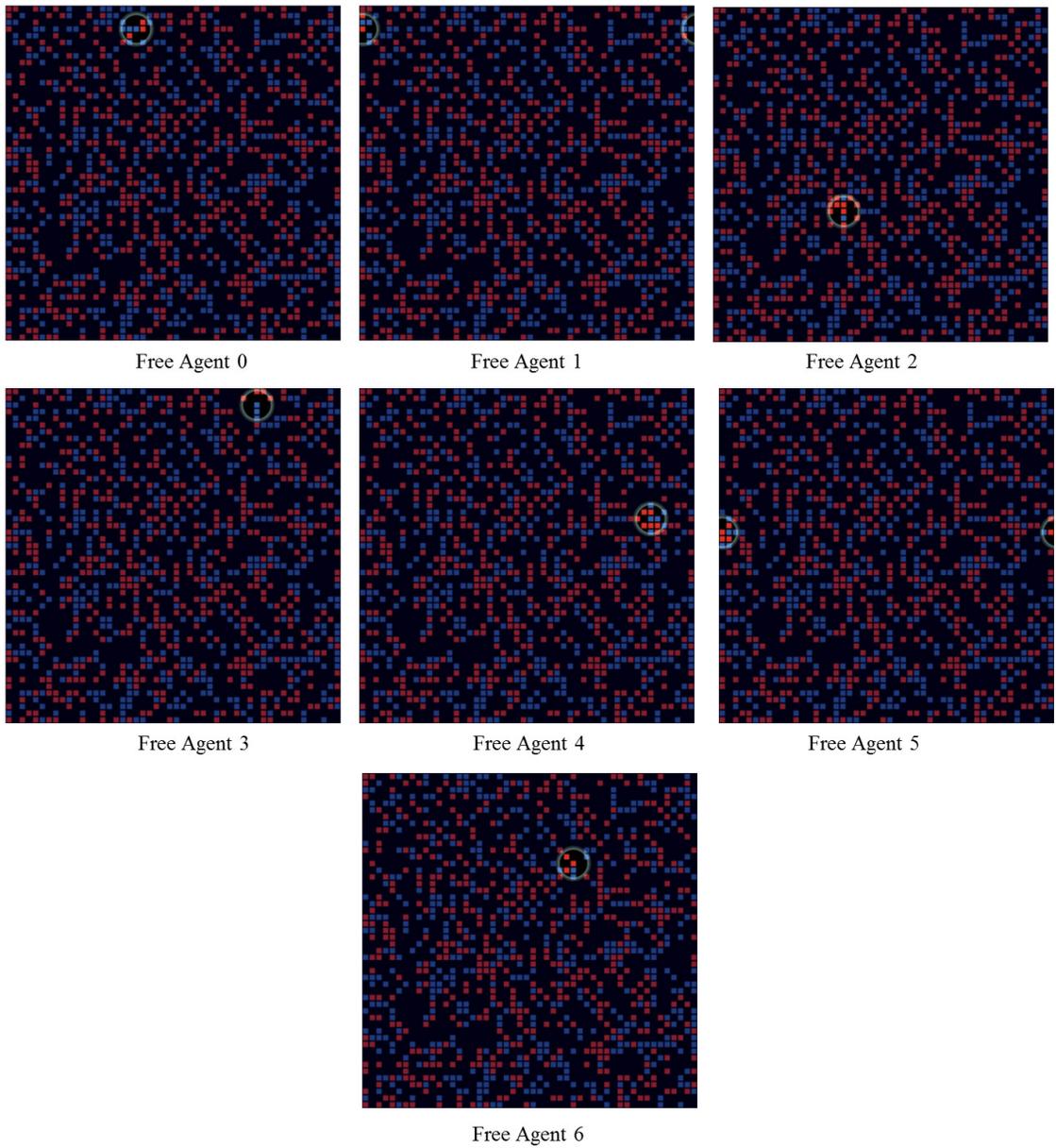


Figure 8-2: Initial Positions of the Free Agents for All Experiments. The images above show the initial position of the free agents employed in each experiment shown in Tables 8-1 and 8-2. The free agents are highlighted in order to make them easier to identify.

Each of the experiments with the normal movement rule produced the identical end state shown in Figure 8-3. After establishing the baseline, runs 1 through 10,000 for each experiment allowed for the designated free agent to make alternative decisions while all other agents behaved identically from run to run.

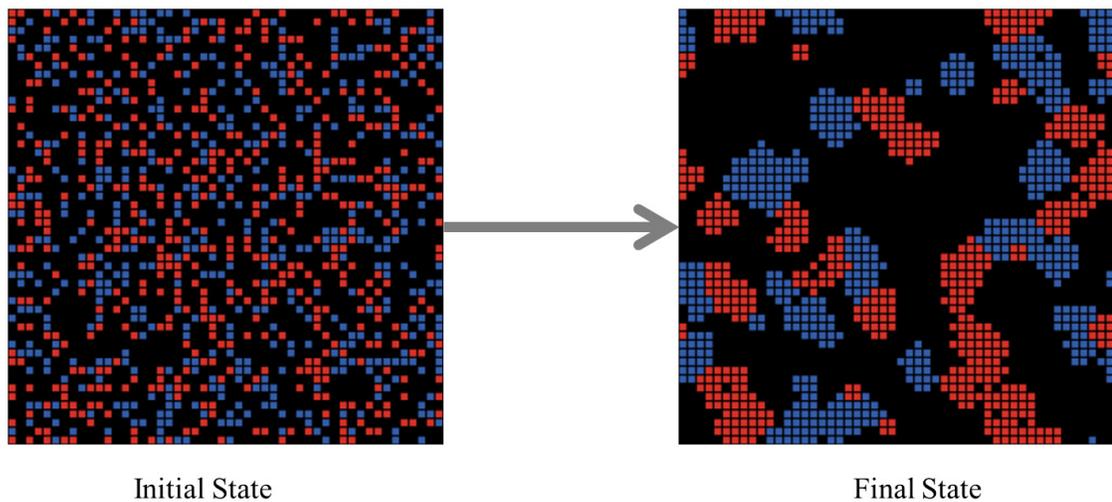


Figure 8-3: Initial and Final State of Run 0 for all Normal Movement Experiments.

In experiments 1.0 through 1.6, the different choices made by each free agent affected the outcomes in different ways. The composite results of all seven experiments, each run 10,000 times in addition to the established baseline, are shown in Figure 8-4.

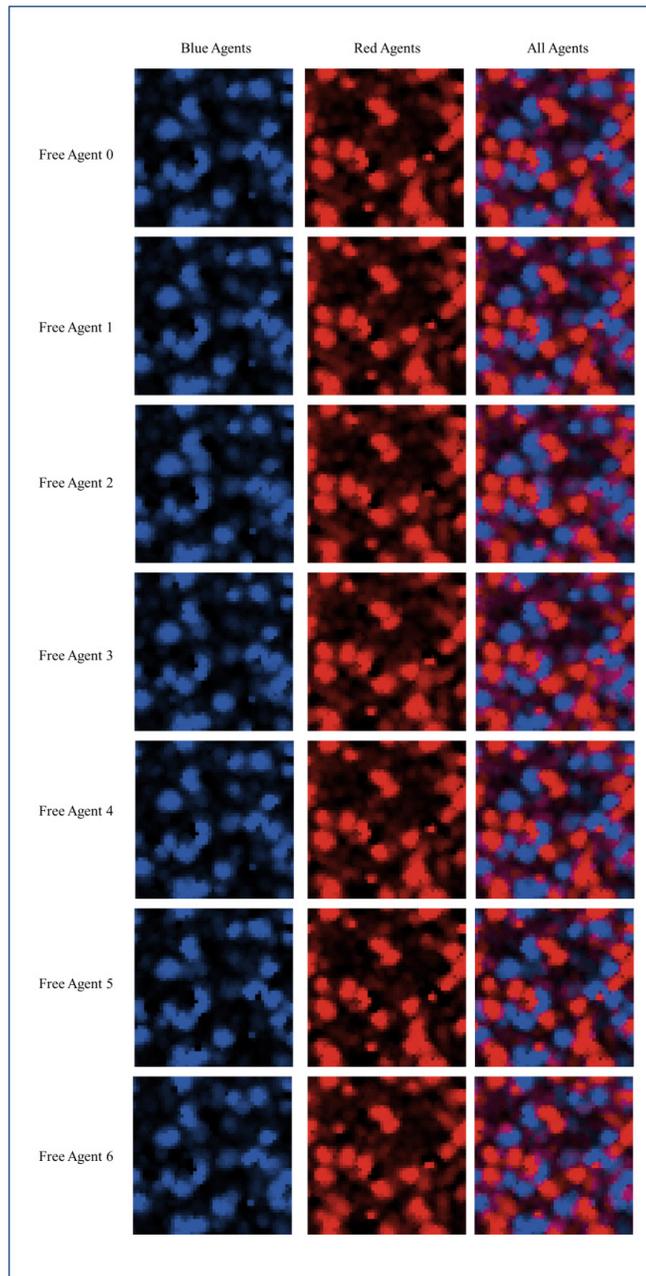


Figure 8-4: Composite Results of Experiments 1.0-1.6. The images constitute composite results of all experiments where agents 0 through 6 were free agents for 10,000 runs each after running the baseline. In each image, the brighter the color, the higher the probability that a red or blue agent settled on that patch at the end of the simulation. The leftmost images show the final probabilities of blue agent locations, the center shows the same for red agent locations, and the rightmost images show both red and blue together where purple areas are those that were either red or blue at the end of the 10,000 runs in the experiment.

Simple visual inspection shows that much of the landscape is structurally consistent across all runs. The centroids of blue and red neighborhoods remain nearly identical under all seven experimental conditions. However, there are also areas where distinct differences occur, predominantly affecting the density of particular results, and the likelihood that certain structural features will become the anchors for more contingent structures. For example, a comparison between the composite results of experiments 1.5 and 1.6, “free agent 5” vs. “free agent 6,” shown in Figure 8-5, reveals how each agent differentially affects the system as a whole.

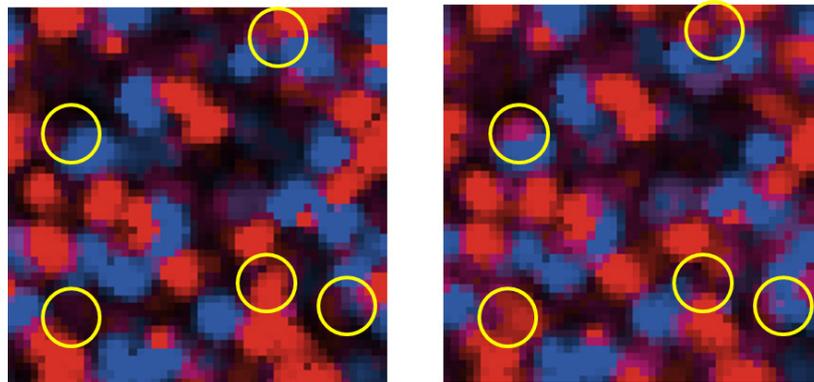


Figure 8-5: Composite Results of Experiment 1.5 and 1.6. These images show the differences between the composite results of experiments 1.5 and 1.6, as generated by agents 5 and 6 respectively. The yellow circles show areas where the composite landscapes differ. Because the only differences between these experiments were the agents making choices, those regions that are different can be attributed to the consequences of agency while those that are common across all cases may be immutable and embedded in the structure of the system itself.

The identified differences show the extended consequences of a single agent making different choices about where to move, and the subsequent cascading consequences of those decisions. While these visual comparisons are useful, they are

incomplete and do not provide any precise measurements of the differences between runs or the extent to which an agent’s choices can produce deviations from the baseline. In order to understand this, the “run distance” metric discussed earlier was calculated for each of the experiments and the distribution of results for each free agent and the 10,000 runs associated with their decision making is shown in Figure 8-6 below.

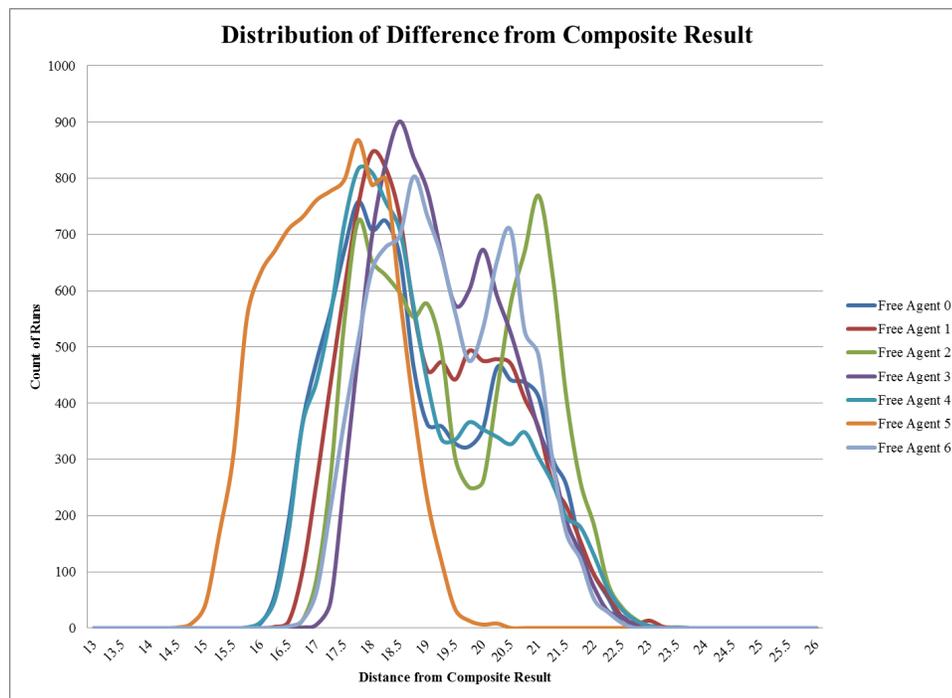


Figure 8-6: The Distribution of the Run Distance Metric for Experiments 1.0-1.6. The distribution of the run distance metric for each of the free agents examined in experiments 1.0-1.6 compares the results of each individual run of the model against the composite results shown in Figure 8-5. The larger the number, the more an individual result varied from the average result depicted by the composite result. The higher the line, the more runs finished with a particular distance from the composite result.

The results in Figure 8-3 are presented again in numerical form in Table 8-3. This table provides the mean, median, min, max, and standard deviation for the results with each free agent.

Table 8-3: Statistical Measures for the Run Distance Metric for Each Free Agent (Normal Movement).

Measure	<i>Free Agent 0</i>	<i>Free Agent 1</i>	<i>Free Agent 2</i>	<i>Free Agent 3</i>	<i>Free Agent 4</i>	<i>Free Agent 5</i>	<i>Free Agent 6</i>
Mean Distance	19.01	19.20	19.63	19.47	18.94	17.35	19.46
Median Distance	18.67	18.95	19.42	19.31	18.60	17.38	19.35
Min Distance	15.99	16.43	16.73	16.99	15.85	14.75	16.59
Max Distance	23.39	23.66	23.05	22.87	23.66	20.36	22.74
Standard Deviation	1.49	1.35	1.44	1.13	1.47	1.01	1.19

The basic statistics and graphical results reveal how each of the examined individual agents affects the system in different ways. Graphically, none of the agents produces a distribution of results that matches the classical normal distribution. Instead, these distributions are largely bimodal, suggesting that once a threshold of differences between the composite and the specific is reached, there is the likelihood that even greater deviations from the composite outcome will result. This bimodal result is most prevalent for free agent 2, experiment 1.3 (shown by the green line in Figure 8-6), and least pronounced for the results for free agent 5 (shown by the orange line in Figure 8-6). More importantly, the results show that free agent 5 had the lowest mean, median distance, minimum, and maximum distance from the composite result, as well as the lowest variance as measured by its standard deviation. Alternatively, free agent 2 had the highest mean and median distance, but agents 1 and 4 produced the greatest maximum value.

These composite measures are largely measures of the distribution of results within each experiment—they do not speak directly to the extent to which free agents can alter the situation as it relates to the baseline result, but do address how much variation their decisions can have on the system as whole. The difference is subtle but important. An agent might have low variation from run to run, but the overall result may be far removed from the configuration established by the baseline. This is possible because of the fact that 10,000 runs of the model may obscure the situation established by one baseline run when averaged together. Therefore, the secondary metric, the percentage difference from the baseline, is needed.

The distribution of the percentage difference from the baseline result, shown in Figure 8-3 earlier, is shown in Figure 8-7 below. Importantly, these results show that free agents can be grouped approximately into two bins, one with a high ability to deviate from the baseline, and the other with far less of an impact on the system. The high-impact agents are 2, 3, and 6, which all produce deviations from the baseline that also approximate a normal distribution—meaning that the extent of the change is relatively predictable even though the specific properties of their spatial results were highly divergent as determined by the variance of the results around the composite outcome. Alternatively, agents 0, 1, 4, and 5 all showed lower deviations from the baseline, even if they possessed multiple modes and larger variance when compared to the more powerful agents.

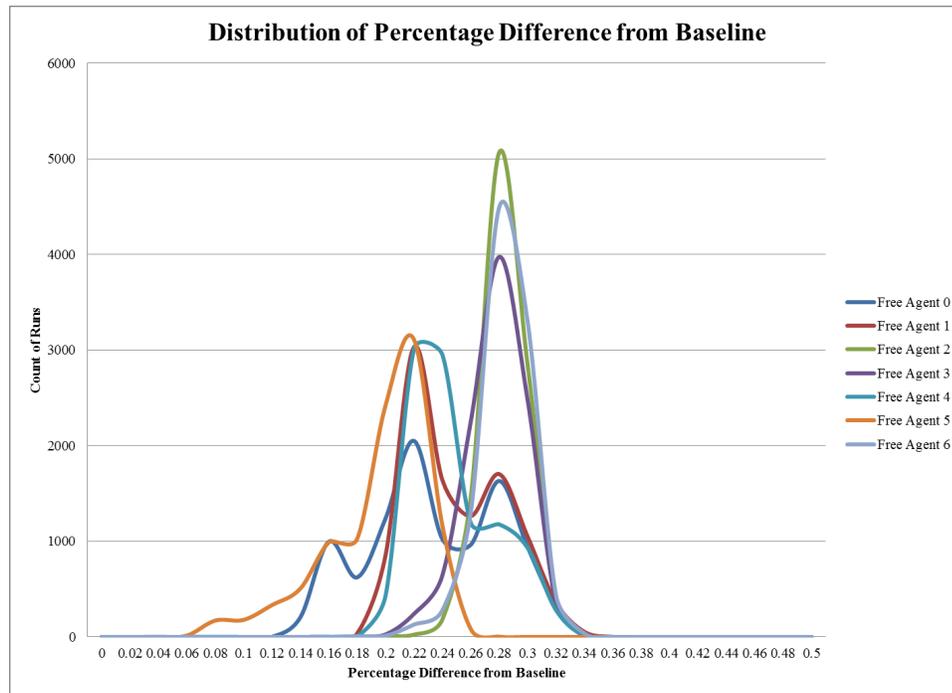


Figure 8-7: The Distribution of the Percentage Difference Metric for 10,000 Runs of Each Free Agent for Experiments 1.0-1.6. The distribution of the percentage difference metric for each of the free agents examined in the experiments 1.0-1.6 compares the results of each individual run of the model against baseline result shown in Figure 8-3. The larger the number, the more an individual result varied from the baseline outcome. The higher the line, the more runs finished with a particular distance from the composite result.

Again, graphically interpreting the results provides an incomplete understanding of the experimental results. Therefore, simple summary statistics on the percentage difference from the baseline are shown in Table 8-4 below.

Table 8-4: Statistical Measures for the Percentage Difference Metric for Each Free Agent (Normal Movement).

Measure	<i>Free Agent 0</i>	<i>Free Agent 1</i>	<i>Free Agent 2</i>	<i>Free Agent 3</i>	<i>Free Agent 4</i>	<i>Free Agent 5</i>	<i>Free Agent 6</i>
Mean Percentage Difference	0.24	0.26	0.29	0.29	0.26	0.21	0.29
Median Percentage Difference	0.24	0.25	0.29	0.29	0.25	0.22	0.30
Min Distance Percentage Difference	0.14	0.19	0.19	0.20	0.09	0.05	0.17
Max Distance Percentage Difference	0.36	0.37	0.35	0.36	0.36	0.28	0.35
Standard Deviation	0.05	0.03	0.02	0.02	0.03	0.04	0.02

These data show that agent 5 produced the smallest percentage difference from the baseline, meaning that it had the least amount of power to shape the system when compared with other agents. Alternatively, agents 2, 3, and 6 each had comparable means with respect to their ability to affect the baseline, but agent 6 had a slightly higher median, suggesting that it had a higher impact on the baseline in slightly more of its 10,000 simulation runs than the other two most influential agents.

Purposeful Movement

The analysis of the model under the conditions of normal movement, i.e. the continuous dependence on the selection of random locations until an unoccupied patch was located for unhappy moving agents, saw significant changes in macroscopic outcomes as a result of individual choices. Indeed, as Table 8-4 noted, the minimum deviation from the baseline result was 5 percent caused by agent 5. However, as suggested earlier, if the agents in the system are willful and act with purpose, then their ability to adapt to others' decisions may limit the power of single agents to influence the system. In order to examine this hypothesis, the same experiments were repeated using the alternative, purposeful movement rule in the model discussed earlier in this chapter.

The first step in evaluating the effects of the purposeful movement rule was to reestablish the simulation baseline for all experiments. This was performed by rerunning the model with a fixed set of random number seeds for all agents starting from the identical initial conditions as the experiments performed in Table 8-1. Thus, the experiments shown in Table 8-2, referred to as experiments 2.0 through 2.6, were identical in every way to those with the normal rule with the only difference being the behavior of the agents when moving.

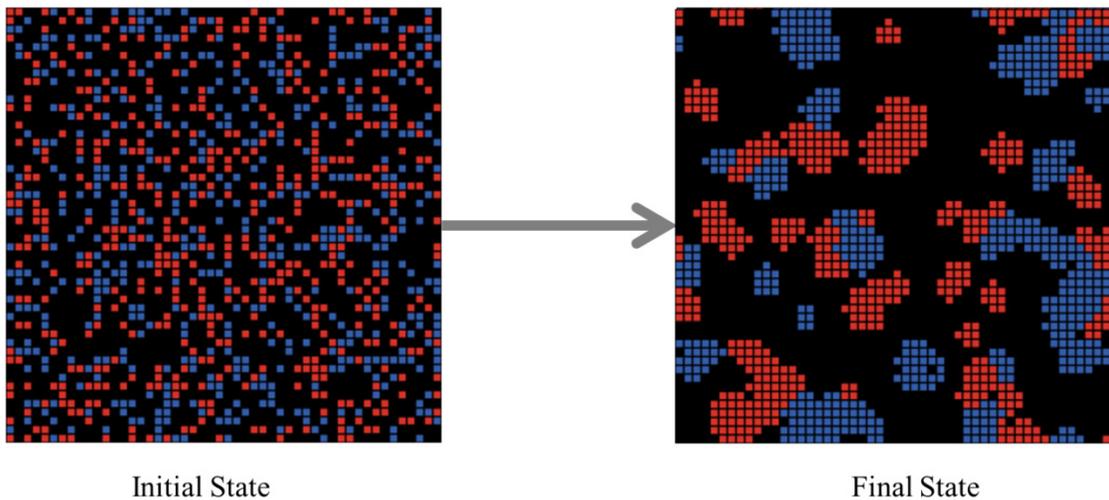


Figure 8-8: The Initial State and Run 0 that Established the Baseline for all Purposeful Movement Experiments.

After establishing the baseline result, the same analytic process was followed. First, composite results of each individual experiment were examined based on different free agents. Afterward, the run distance metrics were examined individually, and also compared with those of experiments 1.0-1.6 in order to compare the extent to which

outcomes varied as a result of change in movement rules. Finally, the percentage difference metrics were examined for experiments 2.0-2.6, comparing the experimental results with the baseline result shown in Figure 8-8 above.

The composite results for experiments 2.0-2.6, characterized by Table 8-2, are shown in Figure 8-9 below. These images show the composite outcomes for 10,000 runs and the established baseline for each of the free agents. The images on the left column show the probabilistic outcomes of where blue agents settled, the center column shows where red agents settled, and finally, the right column shows red and blue together.

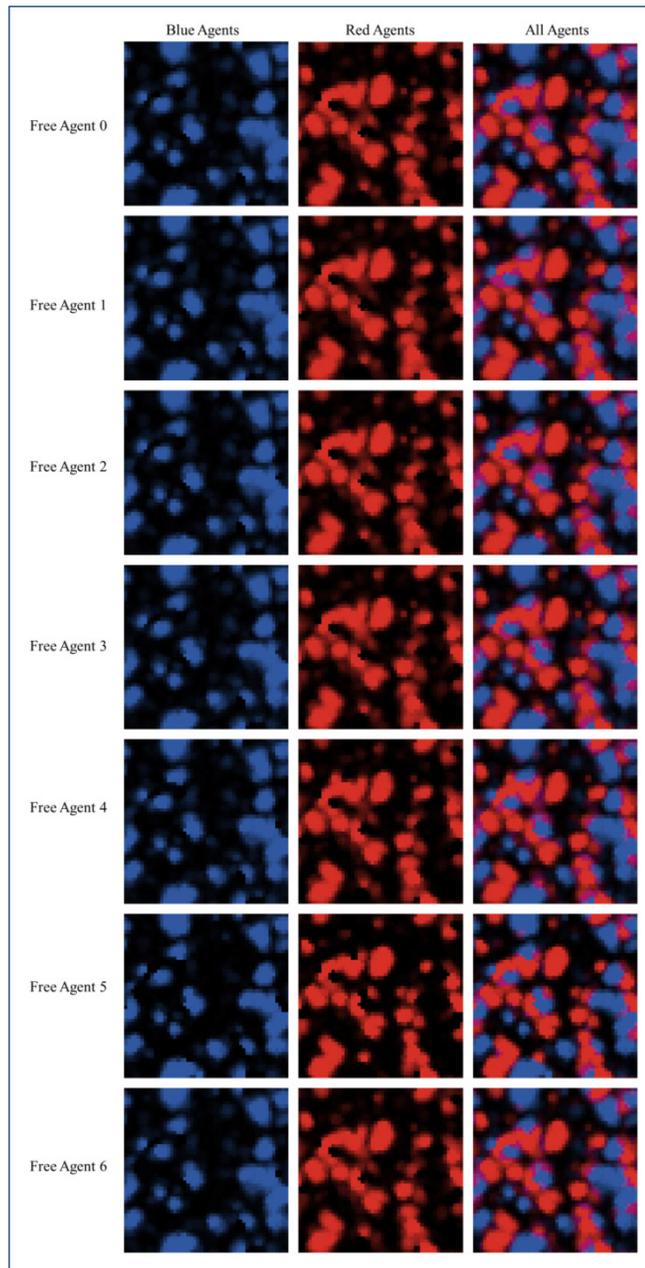


Figure 8-9: Composite Results of Experiments 2.0-2.6. The images constitute composite results of all experiments where agents 0 through 6 were free agents for 10,000 runs each after running the baseline. In each image, the brighter the color, the higher the probability that a red or blue agent settled on that patch at the end of the simulation. The leftmost images show the final probabilities of blue agent locations, the center shows the same for red agent locations, and the rightmost images show both red and blue together where purple areas are those that were either red or blue at the end of the 10,000 runs in the experiment.

As before, the differences between the composite results of each experiment appear small based on visual inspection, again suggesting that much of the landscape is structurally determined. However, as the comparison between free agents 2 and 5 shows, certain contingent results exist that are caused by the differences in a single agent's choices. A comparison between the composite results of agents 2 and 5 is shown in Figure 8-10 below.

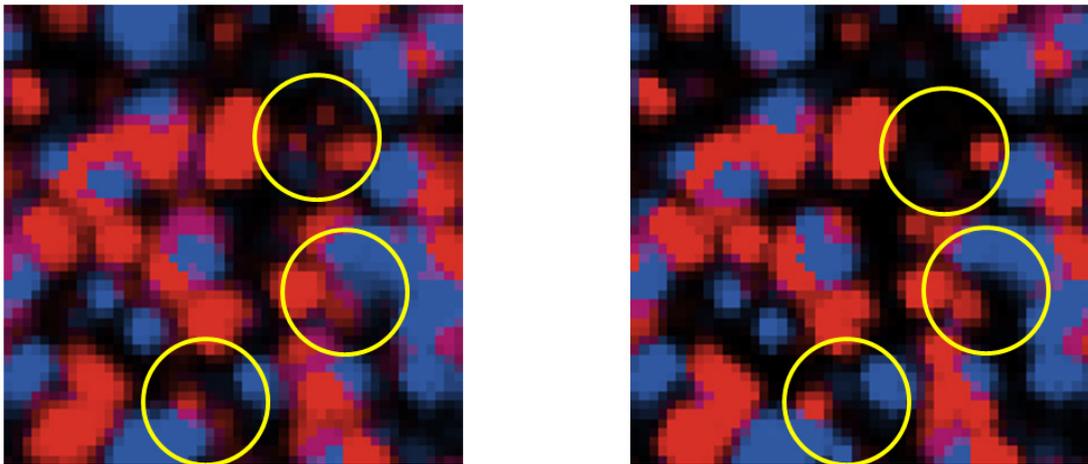


Figure 8-10: Comparison between the Composite Results of Experiments 2.2 and 2.5. The image on the left is the result of experiment 2.2 with free agent 2, and the image on the right is the composite result of experiment 2.5 with free agent 5. The yellow circles highlight areas where the results differ, producing outcomes in one experiment that are not achieved in the other.

As noted earlier, visual inspection alone provides an incomplete assessment of the experimental results and cannot effectively quantify their differences. Therefore, the run

distance metric is shown in Figure 8-11 below, which shows the distribution of differences for each of the experiments and the associated free agents.

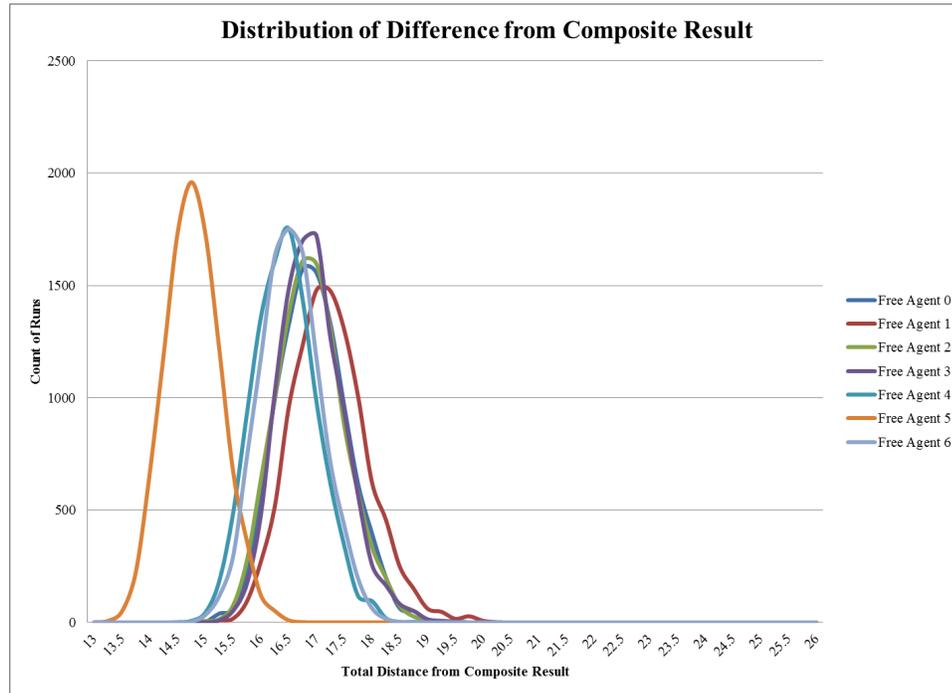


Figure 8-11: Distribution of the Run Difference Metric for Experiments 2.0-2.6.

The differences between the results in Figure 8-11 and the comparable results from the normal movement rule shown in Figure 8-6 are stark. No longer are the results bimodal, as each free agent has produced approximately normal distributions of differences. Moreover, the magnitude of the distances is far smaller, as confirmed by examining the numerical statistics shown in Table 8-5.

Table 8-5: Statistical Measures for the Run Distance Metric for Each Free Agent (Purposeful Movement).

Run Difference Metric	Free Agent 0	Free Agent 1	Free Agent 2	Free Agent 3	Free Agent 4	Free Agent 5	Free Agent 6
Mean Distance	17.06	17.36	17.02	17.05	16.57	14.91	16.66
Median Distance	17.04	17.32	17.00	17.01	16.55	14.90	16.64
Min Distance	15.01	15.34	15.11	15.32	14.73	13.17	14.77
Max Distance	19.62	20.33	19.65	20.09	18.94	16.83	19.58
Standard Deviation	0.62	0.68	0.61	0.59	0.57	0.50	0.55

As Table 8-5 notes, the mean, median, min, and max distances achieved by agents making different decisions all produced fewer deviations from the composite result. Likewise, the standard deviations were approximately half the size of those shown in Table 8-3 under the normal movement rule. Thus, the effects of agency appear significantly muted once the actors in the system are treated as willful and adaptive.

The diminished agency is evident when directly examining the distribution of the distances from their respective composite results between experiments using the same movement rules. Figure 8-12 compares the distributions of the run distance metric for agents 0 and 5 as an illustration of these differences. The results clearly demonstrate the loss of the bimodal structure seen in the normal movement experiments, and much smaller deviations from the composite result.

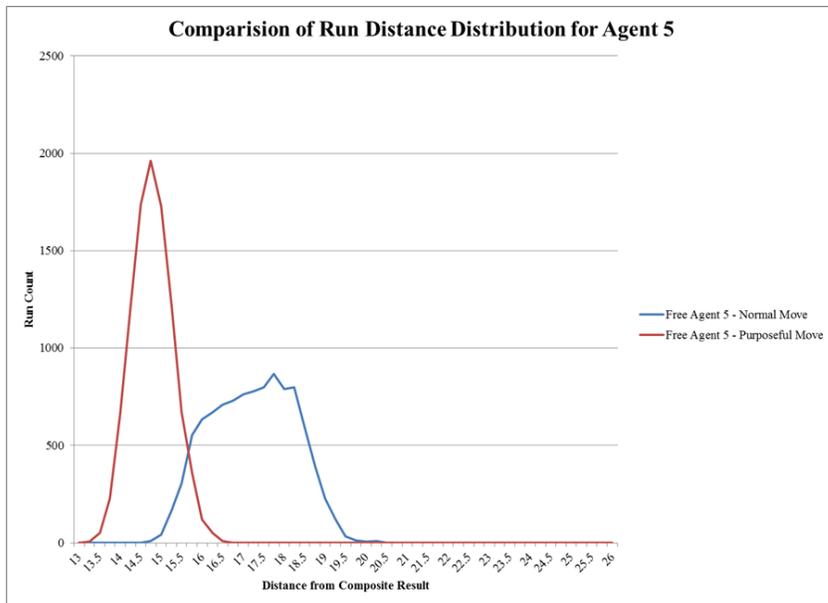
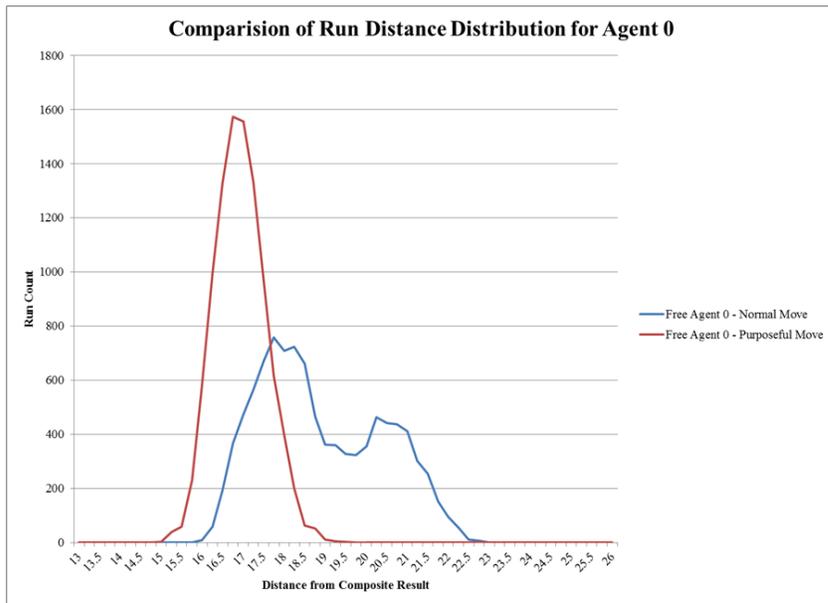


Figure 8-12: Comparison of Run Distance Measure for Free Agents 0 and 5. The top image shows the distribution of the run distance measure, when agent 0 is a free agent, as calculated by comparing individual simulation runs against the composite or average result from that experiment. The blue line shows the distribution under the normal movement rule as originally described in the segregation model, while the red line shows the result from the modified purposeful movement rule. The lower image shows the identical information for agent 5.

The limitations on agency are even more apparent when examining the ways in which purposeful, adaptive actions on the part of the actors in the system limit the extent to which deviations from the baseline occur. The distribution of the percentage difference from the baseline for experiments 2.0-2.6 can be seen in Figure 8-13 below.

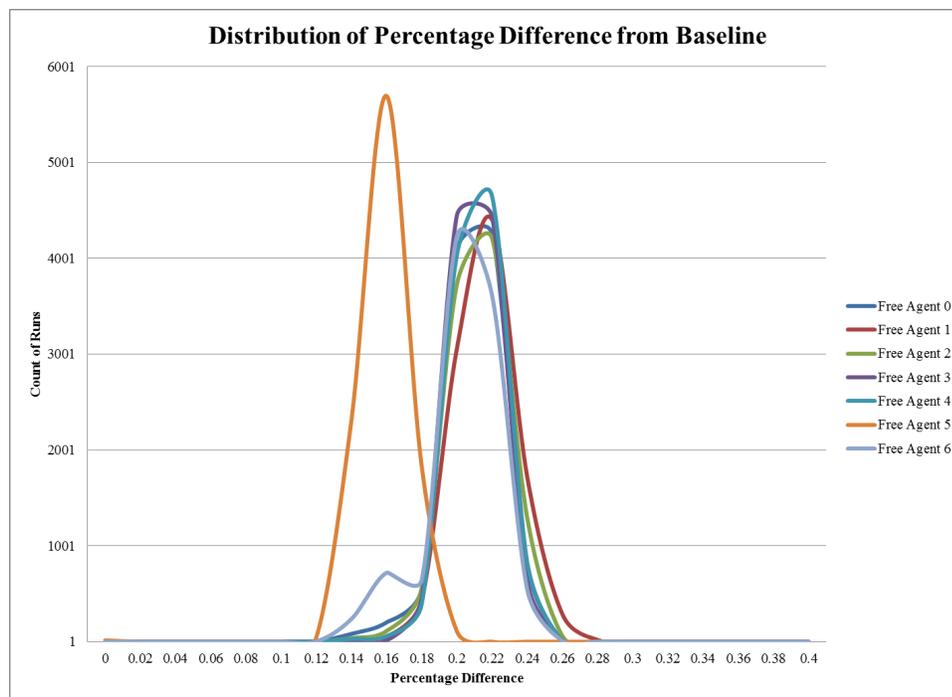


Figure 8-13: The Distribution of the Percentage Difference Metric for 10,000 Runs of Each Free Agent for Experiments 2.0-2.6.

As Figure 8-13 shows, the results are far closer to a normal distribution than those shown in Figure 8-7, showing the same results for the normal movement experiments. Again, the further to the left a line is on the graph, the less a simulation result differed from the

baseline outcome with respect to where red and blue agents settled on the landscape. The extent to which the adaptive actions of the other agents diminished the effects of the isolated free agent in each experiment is even clearer when looking at the numerical results in Table 8-6 below.

Table 8-6: Statistical Measures for the Percentage Difference Metric for Each Free Agent (Purposeful Movement).

Percentage Difference Metric	Free Agent 0	Free Agent 1	Free Agent 2	Free Agent 3	Free Agent 4	Free Agent 5	Free Agent 6
Mean Percentage Difference	0.22	0.23	0.22	0.22	0.22	0.17	0.21
Median Percentage Difference	0.22	0.23	0.22	0.22	0.22	0.17	0.22
Min Distance Percentage Difference	0.13	0.14	0.13	0.15	0.14	0.00	0.14
Max Distance Percentage Difference	0.27	0.29	0.27	0.27	0.27	0.21	0.27
Standard Deviation	0.02	0.02	0.02	0.01	0.01	0.01	0.02

The ranges of outcomes in the purposeful movement experiments were greatly diminished by the adaptive movements of the population. The numerical data shows that the mean and median differences for all agents resulted in lower differences by between 4 and 5 percent, and for agent 6 the greatest reduction of distance was over 8 percent when compared with its relative influence over the system under the normal movement rule. More importantly, agent 5 produced a minimum difference of 0.0 percent fifteen times out of the 10,000 simulations, meaning that no changes from the baseline scenario resulted from making different choices.

Summary

To date, experiments employing ABM have been used to link microlevel decisions and actions with macrolevel structures in order to challenge or confirm the results of macrolevel models, for example mathematical models of the same system using

differential equations. However, by designing and instrumenting models at the level of the individual agent, new experimental and analytic approaches can be developed. First, alternative specifications of ABMs can be examined on the basis of how they achieve balances between agency and structure, encouraging greater research into the characteristics of agency through the development, evaluation, and comparisons of multiple ABMs. Second, ABM can bolster qualitative process-tracing techniques by generating large numbers of cases and scenarios that cannot be constructed out of the historical record alone. Doing so can allow intelligence analysts to identify the relative power of different agents within a system based on differences that result from their choices while holding all other dimensions of the model constant.

The experimental design demonstrated in this chapter is incomplete and only partially addresses the full range of questions and concerns of intelligence analysts. First, as noted earlier, many potential avenues of exploration were not examined, such as changes in the activation order of the agents. Such a change would further open the space of exploration by examining the conditions under which agents are advantaged by acting before others or waiting for others to commit first. The activation order of agents is a particularly interesting property of the model in theoretical terms, because it is not easily categorized as a structural or agent-level variable, but is contingent on the specific model in question.

An additional area of experimentation would extend the extent to which strategic interactions between agents are considered by looking at how combinations of free agents might generate results that might not be identified by examining agents individually. For

example, the demonstration above examined the consequences of agents 0 through 6 making different choices as individual experiments. Additional experiments would explore how combinations of these agents making different choices could reach outcomes that were not achievable under the demonstrated conditions.

These microlevel examinations of ABM fill a critical gap in the existing approaches to intelligence analysis and the ability to support strategy. As noted in Chapter 3, intelligence consumers tend to believe in the power of agency and the importance of their choices, while intelligence producers tend to be more skeptical and reserved about the international system, giving more consideration to structural properties or the adaptive capabilities of other actors that can blunt their consumers' agendas and plans. Using ABM within this context can aid producers and consumers by adjudicating between the two through the use of simulation and experimentation. As a result, many of the most difficult and pressing strategic questions, such as whether decapitating a terrorist organization, developing a new weapon system, attempting to eradicate crops of illicit narcotics, and more can be evaluated in ways that account for the adaptive, creative, and willful decisions of the policy's intended (and equally important, unintended) targets.

Creating the capabilities to fully explore ABMs in the ways advocated for in this chapter is a non-trivial development. It requires considerable theoretical and technical development in order to build the computational infrastructure for handling massive numbers of experiments, the data from simulation results, and the development of appropriate measures of power and influence for characterizing what outcomes can be regarded as systemic or contingent. Moreover, this approach to experimentation allows

for ABMs to examine causation within complex adaptive systems, further aligning the use of models with the needs of microscience and intelligence analysis. By isolating the particular causal paths that result from individual and collective choices, intelligence analysts can better warn of policy makers of potential threats or emerging opportunities.

Chapter 9: Conclusions

The arguments and demonstrations in the previous chapters build toward the conclusion that ABMs provide intelligence analysts and policy makers with new tools for thinking through the identification consequences of choices that are available to themselves and others. These tools may be employed in many ways, such as generating multiple scenarios against which policy makers can prepare themselves and the organizations they lead, to detailed investigations into the strategic dynamics of particular moves and countermoves in order to tease out those aspects of the international system that are structurally determined from those that are contingent. The many uses of models suggest that there is no single outcome waiting for the intelligence community to arrive at regarding the use of ABM in tradecraft. Instead, how analysts, methodologists, managers, collectors, and consumers develop, employ, and infer from ABMs and other modeling formalisms will determine the viability of a model-centric analytic tradecraft.

The Policy Pull

The arguments regarding ABM presented in the preceding chapters have been examined from the perspective of the intelligence community. These discussions have emphasized the “push” from the intelligence community, where intelligence producers may turn to ABM as a new means for providing consumers with better, more complete products. In practice, however, policy makers already demand inputs from ABM and are

“pulling” for analytic products and support that capitalize on the features simulation models rather than rely on analysts’ intuitive judgments alone or mathematical models whose heroic assumptions are too difficult for consumers to accept.

One of the clearest examples of the pull from policy makers for ABMs came from the President of the European Central Bank (ECB), Jean-Claude Trichet. Trichet noted the importance of ABM as a means for addressing deficiencies in traditional macroeconomic policy models, particularly by generating scenarios as a means for augmenting historical experience and in assessing the results of more realistic, boundedly rational behavior in the modeling of complex financial systems. In discussing the inadequacies of the ECB’s analytic capabilities in the midst of the 2008 global financial crisis, he noted:

When the crisis came, the serious limitations of existing economic and financial models immediately became apparent. Arbitrage broke down in many market segments, as markets froze and market participants were gripped by panic. Macro models failed to predict the crisis and seemed incapable of explaining what was happening to the economy in a convincing manner. As a policy-maker during the crisis, I found the available models of limited help. In fact, I would go further: in the face of the crisis, we felt abandoned by conventional tools....

In exercising judgement, we were helped by one area of the economic literature: historical analysis. Historical studies of specific crisis episodes highlighted potential problems which could be expected. And they pointed to possible solutions. Most importantly, the historical record told us what mistakes to avoid....

But relying on judgement inevitably involves risks. We need macroeconomic and financial models to discipline and structure our judgemental analysis. How should such models evolve?...

First, we have to think about how to characterise the homo economicus at the heart of any model. The atomistic, optimising agents underlying existing models do not capture behaviour during a crisis period. We need to deal better with heterogeneity across agents and the interaction among those heterogeneous agents. We need to entertain alternative motivations for economic choices. Behavioural economics

draws on psychology to explain decisions made in crisis circumstances. Agent-based modelling dispenses with the optimisation assumption and allows for more complex interactions between agents. Such approaches are worthy of our attention.

Second, we may need to consider a richer characterisation of expectation formation. Rational expectations theory has brought macroeconomic analysis a long way over the past four decades. But there is a clear need to re-examine this assumption....

Third, we need to better integrate the crucial role played by the financial system into our macroeconomic models. One approach appends a financial sector to the existing framework, but more far-reaching amendments may be required. In particular, dealing with the non-linear behaviour of the financial system will be important, so as to account for the pro-cyclical build up of leverage and vulnerabilities.⁴³⁷

Trichet's comments echoed the key arguments made in previous chapters. First, he noted that the formal mathematical tools failed to provide policy makers with the support they needed. Importantly, the modeling failures identified by Trichet extended beyond their predictive limitations or accuracy. Instead, the fact that these models could not provide policy makers with convincing explanations regarding what was occurring or might happen in the global economy was a critical factor in policy makers' feeling of abandonment.

Second, the insights provided by historical explanation and the study of particular cases proved highly valuable in developing the judgment of policy makers. These lessons were not comprehensive, but their focus on particular cases provided instructive examples regarding how prior policy choices failed to produce desired outcomes. Again, the importance of history was a source of detailed narratives about cause and effect, and the limitations of policy makers' influence.

⁴³⁷ Jean-Claude Trichet, *Reflections on the Nature of Monetary Policy Non-Standard Measures and Finance Theory*, November 19, 2010, available at: <http://www.ecb.int/press/key/date/2010/html/sp101118.en.html> (accessed on October 12, 2012).

Third, the dependence on rational actors and choices is inadequate and misleading. Trichet noted that ABM's ability to model the interactions of boundedly rational, adaptive, learning, and strategic actors was an invaluable contribution in correcting for the heroic assumptions embedded in more traditional macroeconomic models. Indeed, in the context of macroeconomics, ABMs serve as a vehicle for incorporating findings from microeconomics, psychology, sociology, and other fields into the representation of the global economy.

Fourth, analytic models lacked the complexity to realistically cope with interdependent systems in order to generate and capture endogenous and non-linear dynamics. Again, policy makers' need to examine systems of complex interdependent systems requires flexible representations that can capture interactions across scales and boundaries, giving greater emphasis to the synthetic breadth of models rather than their analytic depth. Together, Trichet's points identify why consumers are pulling for the development of a new analytic paradigm that is specifically tailored toward meeting their informational needs.

The experiences of policy makers in the area of international finance and the management of the global economy are not unique. Moreover, they reflect a significant gap between the practice and preferences of social scientists and the needs of policy makers that have been well documented within the international relations and national security community.⁴³⁸ Indeed, national security policy makers and analysts have

⁴³⁸ Alexander L. George, *Bridging the Gap: Theory and Practice in Foreign Policy* (Washington, DC: United States Institute of Peace, 1993); and Miroslav Ninic and Joseph Leggold, eds., *Being Useful: Policy Relevance and International Relations Theory* (Ann Arbor, MI: University of Michigan Press, 2000).

experienced numerous disappointments with conventional analytic tools, methods, and theory. On matters ranging from operations research to social science, the tools and theories available to policy makers have been continuously challenged by the problems of agency, dynamics, and complexity as demonstrated by the schism that occurred during the 1960s between operations research and net assessment over the problems posed assessing long-term trends and developing strategies for coping with dynamic, adaptive, learning competitors.⁴³⁹

Within the domain of national security, real-world policy problems at the tactical, operational, and strategic levels have consistently challenged the available formal analytic tools and leading theories. For example, Stephen Biddle noted that mathematical combat models did not represent the capabilities and behaviors of small, dismounted infantry units that comprised the majority of the Taliban's fighters and military organization, making analysis of the Afghan War problematic.⁴⁴⁰ Likewise, Mark Herman noted conventional, attrition-based measures of military power failed to account for the non-linear effects of degrading command, control, and communications systems; information advantages derived from intelligence sensors and networks; and other

⁴³⁹ Andrew W. Marshall, *Problems of Estimating Military Power* (Santa Monica, CA: RAND, 1966); Eliot A. Cohen, "Net Assessment: An American Approach," *Jaffe Center for Strategic Studies Memorandum No. 29* (Tel Aviv, IL: Tel Aviv University, 1990); Institute for Defense Analysis, *Net Assessment: The Concept, Its Development and Its Future* (Alexandria, VA: Institute for Defense Analysis, May 22, 1990); George E. Pickett, James G. Roche, and Barry D. Watts, "Net Assessment: A Historical Review," in Andrew W. Marshall, J. J. Martin, and Henry S. Rowen, eds., *On Not Confusing Ourselves: Essays on National Security Strategy in Honor of Albert and Roberta Wohlstetter* (Boulder, CO: Westview Press, 1991), pp. 158-185; Stephen Peter Rosen, "Net Assessment as an Analytical Concept," in Andrew W. Marshall, J. J. Martin, and Henry S. Rowen, eds., *On Not Confusing Ourselves: Essays on National Security Strategy in Honor of Albert and Roberta Wohlstetter* (Boulder, CO: Westview Press, 1991), pp. 283-301; Paul Bracken, "Net Assessment: A Practical Guide," *Parameters*, Vol. 36, Spring, pp. 90-100 (2006).

⁴⁴⁰ Stephen Biddle, *Afghanistan and the Future of Warfare: Implications for Army and Defense Policy* (Carlisle, PA: United States Army War College, November 2002), pp. 51-53.

qualitative factors such as training and morale. In each case, these factors characterize agent-level variables that are heterogeneous throughout military organizations.⁴⁴¹

Indeed, these very problems affected the pre-war analysis of Desert Storm in 1991, when analysts estimated that US combat casualties could vary from 3,000 to 45,000, largely based on assumptions that Iraqi military forces were qualitatively similar to Soviet forces employing identical equipment and operational concepts.⁴⁴² However, the actual US combat deaths were far less than even the lowest of predictions, totaling 147 of which 35 were the result of friendly fire—demonstrating the need to reconsider how command control processes, intelligence, communications, training, morale, etc. were handled by analysts and their tools.

From an analytic and theoretical perspective, Desert Storm simultaneously revealed the limitations of existing analytic approaches yet encouraged miscalculations regarding the role of technology in warfare, the utility of military power, and the conflation of operational capabilities and strategic opportunities for shaping the

⁴⁴¹ Mark Herman, “Entropy-Based Warfare: Modeling the Revolution in Military Affairs,” *Joint Forces Quarterly*, No. 20 (Autumn/Winter 1998/99), pp. 85-90.

⁴⁴² Congressional Budget Office, *The Costs of Desert Shield* (Washington, DC: Congressional Budget Office, January 1991), p. 15; and Michael O’Hanlon, “Estimating Casualties in a War to Overthrow Saddam,” *Orbis* (Winter 2003), pp. 24-27.

international system.⁴⁴³ Projecting the 1991 conflict as a template for the future of armed conflict and the use of force in the international system replicated the very tendencies that made the Cold War's sudden conclusion so surprising: a reliance on theories and methods that were overly committed to state-centric models of power and stability that paid little attention to the internal health and vitality of domestic political institutions and factors, e.g. legitimacy and ideology.⁴⁴⁴

The brief period between the Cold War's conclusion and Desert Storm resulted in two notable developments that further challenged analysts and analytic support to policy makers. First, the analytic difficulties associated with understanding emerging military capabilities and opportunities stimulated a demand for new tools and methods that could credibly address agent-level variables, such as information and training, and also generate alternative futures against which different policy choices could be compared. Thus,

⁴⁴³ William J. Perry, "Desert Storm and Deterrence in the Future," in Joseph S. Nye, Jr. and Roger K. Smith, eds., *After the Storm: Lessons From the Gulf War* (Lanham, MD: Madison Books, 1992), pp. 241-263; Alan D. Campen, ed., *The First Information War* (Fairfax, VA: AFCEA International Press, 1992); William A. Owens, "JROC: Harnessing the Revolution in Military Affairs," *Joint Forces Quarterly*, No. 5 (Summer 1994), pp. 55-57; Thomas A. Keaney and Eliot A. Cohen, *Revolution in Warfare? Air Power in the Persian Gulf* (Annapolis, MD: Naval Institute Press, 1995); Joseph S. Nye, Jr. and William A. Owens, "America's Information Edge," *Foreign Affairs*, Vol. 75, No. 2 (March/April 1996), pp. 20-36; Stephen Biddle, "Victory Misunderstood: What the Gulf War Tells Us about the Future of Conflict," *International Security*, Vol. 21, No. 2 (Autumn 1996), pp. 139-179; Bill Owens, with Ed Offley, *Lifting the Fog of War* (New York, NY: Farrar, Straus and Giroux, 2000); David S. Alberts, John J. Gartska, Richard E. Hayes, and David A. Signori, *Understanding Information Age Warfare* (Washington, D.C.: DOD Command and Control Research Project, 2001); Andrew F. Krepinevich, Jr., *The Military-Technical Revolution: A Preliminary Assessment* (Washington, DC: Center for Strategic and Budgetary Assessments, 2002); and James Mann, *Rise of the Vulcans: The History of Bush's War Cabinet* (New York, NY: Viking, 2004), pp. 198-215

⁴⁴⁴ John Lewis Gaddis, *The United States and the End of the Cold War: Implications, Reconsiderations, Provocations* (New York, NY: Oxford University Press, 1994); Richard Ned Lebow and Thomas Risse-Kappen, eds., *International Relations Theory at the End of the Cold War* (New York, NY: Columbia University Press, 1994); Michael W. Doyle and G. John Ikenberry, eds., *New Thinking in International Relations Theory* (Boulder, CO: Westview Press, 1997); John Lewis Gaddis, *We Now Know: Rethinking Cold War History* (New York, NY: Oxford University Press, 1998), pp. 283-284; and Ethan B. Kapstein and Michael Mastanduno, eds., *Unipolar Politics: Realism and State Strategies After the Cold War* (New York, NY: Columbia University Press, 1999).

much of the analytic community, particularly those in the orbit of military planning and intelligence, sought the capabilities that ABMs provide in recognition that their conventional tools were not well suited to answer the questions policy makers were asking.⁴⁴⁵ Perhaps the most explicit acknowledgement of the need for new analytic approaches within the military was made by the Transformation Study Group in 2001, which advised the Secretary of Defense to broaden the mission of the Joint Warfare Analysis Center (JWAC) to understand and model adversaries as complex adaptive systems.⁴⁴⁶

The second development within the area of national security occurred at the intersection between theory and policy. New national security issues including the environment, state failure, terrorism, crime, and cyberspace challenged the traditional realist and liberal emphasis on great powers and national states in thinking about the structure and dynamics of the international system. Thus, the study of arms races, military balances, nuclear deterrence, and the rationalizing tendencies imposed by bureaucracy and markets in service of the state were joined by concerns over transnational networks and flows of goods, wealth, weapons, and people; the diffusion of ideology and technology; and other social processes that emphasized change and

⁴⁴⁵ Desmond Saunders-Newton, "Computational Social Science, Operations Research, and Effects-Based Operations: The Challenge of Inferring 'Effects' from Dynamic Socio-Physical Systems," *Analyzing Effects-Based Operations*, Military Operations Research Society, McLean, VA, January 28, 2002; and Desmond Saunders-Newton and Aaron B. Frank, "Effects-Based Operations: Building the Analytic Tools," *Defense Horizons* No. 19 (October 2002).

⁴⁴⁶ Transformation Study Group, *Transformation Study Report: Executive Summary* (Washington, DC: Department of Defense, April 27, 2001), p. 11; and Transformation Study Group, *Transformation Study Report: Transforming Military Operational Capabilities* (Washington, DC: Department of Defense, April 27, 2001), p. 28.

dynamics over stability and continuity within the international system.⁴⁴⁷ The result was a “renaissance” in security studies that tilted the balance within international relations theory away from realism and liberalism toward constructivism, historical sociology, and the English School as ideas, beliefs, and social communities.⁴⁴⁸ Against this backdrop, the need for analytic capabilities such as those provided by ABM grew while policy makers and policy-oriented scholars saw established international relations theory and research methods as irrelevant to their analytic needs.⁴⁴⁹

The Problem of Social Identity

The relationships between intelligence, policy, and social science are evident in concerns over identity and its role in national security. Social identity has posed a persistent challenge within international relations and the social sciences in general, and is one of the factors in the international system that has come to the forefront of contemporary security considerations. The emphasis on states, rationality, and system

⁴⁴⁷ Kenneth N. Waltz, *Theory of International Politics* (Reading, MA: McGraw-Hill, 1979); Hans J. Morganthau, *Politics Among Nations: The Struggle for Power and Peace* (Boston, MA: McGraw-Hill, 1993); John J. Mearsheimer, *The Tragedy of Great Power Politics* (New York, NY: W. W. Norton and Company, 2001); G. John Ikenberry, *After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order After Major Wars* (Princeton, NJ: Princeton University Press, 2001); and interview with Carmen Medina, Deloitte, July 3, 2012.

⁴⁴⁸ Helga Haftendorn, “The Security Puzzle: Theory-Building and Discipline-Building in International Security,” *International Studies Quarterly*, Vol. 35, No. 1 (March 1991), pp. 3-17; Stephen M. Walt, “The Renaissance of Security Studies,” *International Studies Quarterly*, Vol. 35, No. 2 (June 1991), pp. 211-239; Alexander Wendt, *Social Theory of International Politics* (New York, NY: Cambridge University Press, 1999); Doug McAdam, Sidney Tarrow, and Charles Tilly, *Dynamics of Contention* (New York, NY: Cambridge University Press, 2001); Headley Bull, *The Anarchical Society: A Study of Order in World Politics* (New York, NY: Columbia University Press, 2002); and Shibley Telhami and Michael Barnett, eds., *Identity and Foreign Policy in the Middle East* (New York, NY: Cornell University Press, 2002).

⁴⁴⁹ Ian S. Lustick, Dan Miodownik, and Roy J. Eidelson, “Secessionism in Multicultural States: Does Sharing Power Prevent or Encourage It?” *American Political Science Review*, Vol. 98, No. 2 (May 2004), pp. 211-212; and David Kilcullen, *The Accidental Guerrilla: Fighting Small Wars in the Midst of a Big One* (New York, NY: Oxford University Press, 2009), p. 296.

structure all militate against considerations of actors' identities, how they are formed, change, persist, and influence decision making.

Structural approaches to international relations argue that states are undifferentiated units that only vary with respect to their relative power.⁴⁵⁰ This "third image" approach locates the sources of international conflict within the structure of the international system itself, rather than in the character of the states in the system, e.g. democratic or fascist, or in human nature. Ironically, the absence of identity as a variable in the international system was something that Keohane's seminal theory of regimes confronted, and ultimately argued provided the means for limiting or even escaping the problem of anarchy in the international system.

What distinguishes my argument from structural Realism is my emphasis on the effects of international institutions and practices on state behavior. The distribution of power, stressed by Realists, is surely important. So is the distribution of wealth. But human activity at the international level also exerts significant effects. International regimes alter the information available to governments and the opportunities open to them; commitments made to support such institutions can only be broken at a cost to reputation. International regimes therefore change the calculations of advantage that governments make. To try to understand state behavior simply by combining the structural Realist theory based on distribution of power and wealth with the foreign policy analyst's stress on choice, without understanding international regimes, would be like trying to account for competition and collusion among oligopolistic business firms without bothering to ascertain whether their leaders met together regularly, whether they belonged to the same trade associations, or whether they had developed informal means of coordinating behavior without direct communication.⁴⁵¹

⁴⁵⁰ Kenneth N. Waltz, *Theory of International Politics* (Reading, MA: McGraw Hill, 1979), pp. 97, and 143-144.

⁴⁵¹ Robert O. Keohane, *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton, NJ: Princeton University Press, 2005), p. 26.

According to Keohane, regimes could ultimately normalize, or socialize, relations between states by introducing notions of history and individual identity into the international system by developing reputations for states that established which states honored their agreements and operated within accepted norms, and those that did not. Thus, identity's appearance in international relations subtly challenged structural perspectives by demanding careful attention to agents' history, learning, and information.

The importance of identity in international relations and national security was made even more explicit by Philip Bobbitt, who argued that identities bound domestic constitutional politics with international relations and strategy. Bobbitt's approach to international relations was characterized by the interdependent relationship between constitutional law and international strategy—where each selected the other, creating an evolutionary dynamic in which the success and failure of states simultaneously depended on their internal organization and relations with others. At the center of his theory was identity, which, according to Bobbitt, was the state's historical experience and narrative, i.e. society's self-image that bounded acceptable constitutional structures and international strategies.

Law, strategy, history—three ancient ideas whose interrelationship was perhaps far clearer to the ancients than it is to us, for we are inclined to treat these subjects as separate modern disciplines. Within each subject we expect economic or political or perhaps sociological causes to account for developments; we are unlikely to see any necessary relation among these three classical ideas. They do not appear to depend upon each other....

The State exists by virtue of its purposes, and among these are a drive for survival and freedom of action, which is strategy; for authority and legitimacy, which is law; for identity, which is history. To put it differently, there is no state without strategy, law, and history, and, to complicate matters, these three are not merely interrelated elements, they

are elements each composed at least partly of the others. The precise nature of this composition defines a particular state and is the result of many choices. States may be militaristic, legalistic, and traditional to varying degrees, but every state is some combination of these elements and can be contrasted with every other state—and with its own predecessors—in these ways.⁴⁵²

Bobbitt's approach to international relations introduced many of the concerns that ABMs were developed to address. Because Bobbitt viewed the international system as evolutionary and historical, the sequence in which events occurred mattered, meaning that history feeds back into the behavior of actors, places actors and events in the context of time and place, while preventing the reduction of the international system into a set of independently assessable cases or data.⁴⁵³ Moreover, by linking history, identity, and state formation together, Bobbitt erased the boundaries between domestic and international politics that international relations theorists, particular realists, had dismissed in order to differentiate international relations from the rest of political science.⁴⁵⁴ Therefore, Bobbitt, like others, largely dismissed significant portions of international relations theory in favor interdisciplinary approaches that were better able provide insights into political, military, and legal systems and dynamics across spatial, governing, and temporal scales.

While Keohane and Bobbitt showed the increasing inclusion of identity into theories of international relations and national security, applied social scientists and

⁴⁵² Philip Bobbitt, *The Shield of Achilles: War, Peace, and the Course of History* (New York, NY: Vintage, 2002), pp. 5-6.

⁴⁵³ Richard E. Neustadt and Ernst R. May, *Thinking in Time: The Uses of History for Decision Makers* (New York, NY: The Free Press, 1986).

⁴⁵⁴ Torbjorn L. Knusten, *A History of International Relations Theory* (New York, NY: Manchester University Press, 1997); and James E. Dougherty and Robert L. Pfaltzgraff, Jr., *Contenting Theories of International Relations: A Comprehensive Survey* (New York, NY: Pearson, 2009).

policy makers matched developing theory with empirical research on the role of social identity in particular cases. Multiple studies determined social identities (or more precisely the performance of identities) to be a source of political and social conflict and mobilization, at the domestic, international, and regional levels. The range of cases included contested relations between Hui Muslims in China's Northwest and the central government, the radicalization of Palestinian youth in Lebanese refugee camps, regional concerns over the potential for a pan-Shia identity movement—the Shia Revival—and the Post-Cold War “Clash of Civilizations.”⁴⁵⁵ Likewise, Jerold Post and Mark Sageman each noted the importance of identity in counterterrorism by describing how identity roles bound individuals to larger groups that developed and sustained social commitments and a willingness to sacrifice, lead, and follow.⁴⁵⁶ Finally, social identities have been seen as the outcome of successful state formation and building projects, which establish the legitimacy and authority of government and stabilize societies by giving structure to patterns of social, political, and economic life.⁴⁵⁷

As identity has become increasingly recognized as a contributing factor in multiple national security and intelligence issues, it has become a fertile ground for ABM. This is due to the fact that identities exist simultaneously within agents and at larger social or institutional levels and become operative through the decision making and

⁴⁵⁵ Samuel P. Huntington, *The Clash of Civilizations and the Remaking of World Order* (New York, NY: Touchstone, 1996); Dale F. Eickelman and James Piscatori, *Muslim Politics* (Princeton, NJ: Princeton University Press, 2004); Vali Nasr, *The Shia Revival* (New York, NY: W. W. Norton and Company, 2007); and Bernard Rougier, *Everyday Jihad* (Cambridge, MA: Harvard University Press, 2007).

⁴⁵⁶ Jerrold M. Post, *The Mind of the Terrorist: The Psychology of Terrorism from the IRA to Al-Qaeda* (New York, NY: Palgrave, 2007); and Marc Sageman, *Understanding Terror Networks* (Philadelphia, PA: University of Pennsylvania Press, 2004).

⁴⁵⁷ Ashraf Ghani and Clare Lockhart, *Fixing Failed States: A Framework for Rebuilding a Fractured World* (New York, NY: Oxford University Press, 2009), p. 132.

actions of individuals and groups that define themselves through the comparison of similarities and differences with others. Identities are created dynamically, and are both selected endogenously by agents, the act of identification, and imposed exogenously by others, i.e. the act of being categorized. They provide heuristics for social behavior, problem solving, and determine acceptable and unacceptable beliefs and prescriptions for action based on context. As a result, identities simultaneously serve as the basis for the bounded rationality of individuals and groups, and also provide an alternative to utilitarian, cost-benefit explanations of decision making by offering deontic elements—agents’ rights, privileges, obligations, commitments, rituals, customs, etc.—as explanations for decisions and actions.⁴⁵⁸ When viewed through the lens of identity theory, the challenges that analysts and policy makers face in understanding and influencing the behavior of others become clearer, as do many of the sources of their errors, e.g. mirror-imaging.

Developing a Model-Centric Analytic Tradecraft

The challenge posed by social identity exemplifies many of the problems analysts and policy makers confront. Indeed, efforts to incorporate social identity into national security planning and policy making show the difficulties posed by complex adaptive systems. Identities are too complex to examine on strictly intuitive grounds, but poorly addressed by mathematical models due to the fact that local interactions between

⁴⁵⁸ Steph Lawler, *Identity: Sociological Perspectives* (Malden, MA: Polity, 2008); Richard Jenkins, *Social Identity* (New York, NY: Routledge, 2008); Peter J. Burke and Jan E. Stets, *Identity Theory* (New York, NY, Oxford University Press, 2009); and George A. Akerlof and Rachel E. Kranton, *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being* (Princeton, NJ, Princeton University Press, 2010).

individuals performing, interpreting, and validating roles creates perpetual novelty and innovation. Yet, understanding how they work is crucial for countering radicalization, demobilizing terrorist and insurgent movements, and developing international institutions on matters as diverse as the environment, the global economy and trade, and military weapons and alliances. As a result, ABMs may serve as important analytic tools for overcoming cognitive biases, such as mirror-imaging or rationality, while providing the necessary flexibility to represent intelligence targets as bounded rational, adaptive actors with their own unique decision-making heuristics and social and historical contexts. In these cases, ABMs provide opportunities to improve the quality of empirical models of complex adaptive systems, thus bridging gaps between the intelligence and scientific communities whenever common problems are confronted.

The foremost challenges of analytic tradecraft involve building relations of trust and respect with policy makers so that consumers—who possess their own expertise, worldviews, and sources of information—accept intelligence assessments as authoritative, timely, and relevant to their needs. This is most important in cases when intelligence assessments challenge the deep-seated beliefs of consumers and the assessments of their policy staff regarding the efficacy or consequences of proposed policy actions (or inaction), or when producers and consumers possess divergent expectations about future threats and opportunities. It is this relationship between intelligence producers and consumers, first and foremost, that ultimately sets the parameters for selecting analytic methods.

Prior intelligence studies have examined the producer/consumer relationship and argued that navigating these troubled waters necessarily requires analysts view their work as an art rather than science. Alternatively, many in the intelligence community have noted that scientific approaches to analysis provide the best opportunities for mitigating cognitive biases or mindsets, and provide the necessary self-corrective basis of building knowledge about an uncertain, dynamic, and deceptive international system. The employment of ABM as a core capability in a model-centric tradecraft provides an opportunity to transform this long-standing debate over intelligence as an art or science, tipping the balance toward science through the use of model building as a creative and trust-building activity focused on exposing consumers to the logical consequences of their assumptions before suggesting alternative and challenging perspectives. ABM provides the benefits of formal modeling methods without making heroic assumptions regarding the behavior of a system's units, e.g. perfect rationality or spatial mixing, while remaining flexible enough to represent multiple, competing worldviews or system descriptions in order to explore the consequences of alternative assumptions.

A model-centric analytic tradecraft is simultaneously scientific and normative. It is scientific in the sense that it promotes structured, transparent, and testable relations between evidence and inference—building upon the contemporary development and use of SATs within the intelligence community. It is normative in the sense that the primary objective is to assist consumers in decision making by providing new ways of asking and answering questions about counterfactuals, scenarios, and agency. This allows for the development of a microscience that systematically and logically searches for challenging

cases and explores prospective policy options that reside beyond the limitations of empirical analysis, and do so in a way that is accessible to consumers. This combination of design flexibility and the ability to generate and explore cases beyond the limits of the empirical record adds additional rigor to the existing suite of SATs, rendering their static, often graphical models of intelligence targets into dynamic computational models that extend and formalize subject matter expertise by drawing upon analysts' experience, policy makers' worldviews, and social science theory.

Because intelligence analysis is fundamentally a normative activity conducted with the specific purpose of assisting policy makers, it differs from empirical science with respect to its objectives and has more in common with engineering, law, architecture, medicine, and other artificial or design sciences. Indeed, while philosophers and scientists may disagree as to how science works or should be defined, a near consensus exists regarding its proper emphasis on empiricism, induction, and generalization regarding the desired epistemic properties of scientific outputs. In service to policy makers, however, these epistemic properties have proven problematic when confronted by consumers' needs for information about particular cases and the potential for different outcomes as a result of the strategic choices and actions of actors.

Rather than pursue generalized knowledge that applies to populations of cases, and therefore favors the analysis of structural variables or universal logic, ABMs deployed in a normative role in intelligence tradecraft can enable analysts to focus on agency and particular cases. This means focusing on how small changes in decision making, actions, and interactions can alter the trajectory of systems. Likewise, it means

modeling single cases and systems from multiple perspectives in order to explore the implications of alternative assumptions, theories, and data sources. Thus, analysts can use models as a means for relationship building with consumers and not simply as devices for predicting or exploring the behavior of target systems. By building representations of consumers' and analysts' mental models, they can be made aware of the logical consequences of their assumptions (within the bounds of representational and implementation choices) that are too complex to tease out intuitively. Once gaps between expectations and outcomes are identified, models can be extended to include additional factors or adapted to change assumptions, thus serving as a platform for facilitating a dialog about the international system between producers and consumers.

Future Research

The development of a model-centric analytic tradecraft for the intelligence community remains incomplete. Unresolved questions and needed capabilities span a variety of theoretical, philosophical, and technological grounds, and their resolution requires an interdisciplinary research and development effort. These research needs span the gamut from resolving difficult questions of epistemology, counterfactuals, experimental design, and model implementation and instrumentation, to more general developments within the social sciences that emphasize agency over structure. In each of these instances, a series of interrelated questions exists regarding how to develop the knowledge and expertise to implement a model-centric analytic tradecraft.

The first area for continued research is not unique to the intelligence community or analytic tradecraft. However, the continued development of computational social

science and agency-oriented models of social behavior are needed in order to develop theories about how individual actions and decisions aggregate to affect entire systems (or why in some cases individual decisions have no effect on systems) in order to counter an overreliance on structural explanations and perfect rationality found in the social sciences. The continued development and maturation of ABM as a research tool for investigating social phenomena can provide a renewed social science foundation that is consistent with policy makers' philosophical beliefs about agency. Thus, intelligence analysts and consumers stand to gain significantly as computational social science continues to develop.

Chapters 3 and 4 discussed the particular needs of intelligence analysts and policy makers with respect to agency, counterfactuals, and scenarios. In addressing these questions, simulation, broad search, and experimental design are crucial research components that do not fit within the established disciplinary paradigms of the social sciences that rely upon large- N statistical analyses, small- N case studies, or formal mathematical models employing heroic assumptions. Although ABMs afford scholars and intelligence analysts alike with new research tools, precisely how to employ them scientifically remains an unresolved question. Here, a second line of interdisciplinary research that bridges computer science, social science, and the philosophy of science would fill a notable void. For example, while the literature employing ABM as a research method has proliferated significantly, no specific textbook exists regarding experimental design with multiagent systems. Instead, researchers have relied on designs developed for the use of statistical or mathematical models, which possess far fewer

dimensions and potential modes of behavior. Thus, making microscience a reality will require rethinking the philosophy of science regarding structure and agency, and delineating the strengths and weaknesses of deriving knowledge from ABMs.

A third line of continued research would be predominately technical and ensure that computational models are designed in ways that can align with the informational needs of intelligence consumers and the philosophers who have thought through the epistemology of agency. Chapter 8 demonstrated the potential of ABM to explore the agent-structure relationship in the international system. However, even the simple demonstration was challenged by the limitations of popular ABM development platforms, e.g. NetLogo. Likewise, Chapter 7's demonstration of simulating intelligence collection revealed the importance of model instrumentation.

Designing ABMs that can meet new and complex experimental designs requires new ways of capturing and analyzing data, as well as controlling input parameters in order to search through high-dimensional spaces requiring a thousand, million, or more simulation runs. Future data analysis should pay equal attention to the time series of microlevel, inter-agent interactions in order to capture and identify causal paths within artificial societies. Additionally, efforts should proceed to examine the ways in which models can be instrumented in ways that replicate real-world information collection in order to better understand intelligence collection strategies and capabilities, as well as uncover how collection sources and methods may possess inherent biases and limitations that analysts must be familiar with.

A fourth area of technical research is in the capturing and translation of mental models into computational ABMs. Because large numbers of intelligence analysts and consumers are unlikely to possess rich modeling expertise or the time to spend with modeling, devising methods for rapidly capturing how experts characterize complex systems and encoding their descriptions into software is a necessity if ABMs are going to transition into widespread professional use. Research of this kind would largely involve the design of model interfaces, graphical model-building tools, and development of tools for rapidly translating subject matter expertise into model specifications that can be coded as algorithms.

Finally, these lines of research must come together to support the development and use of ABMs in the production of finished intelligence products. Regardless of developments in social science, philosophy, and technology, only the successful use of ABM by the intelligence community can provide the institutional motivation for its continued use and development in support of consumers. Thus, many of the requirements for successfully developing a model-centric tradecraft cannot be met through pure research alone, but will require successful application—having analysts, methodologists, and outside experts work together to develop intelligence products that meet the needs of consumers in ways that other methods cannot. Until this occurs, sustained institutional pressures for transforming analytic tradecraft may be inconsistent and uneven.

Appendix

Interview with Leon Fuerth, September 21, 2011

Background: On September 21, 2011 I was fortunate enough to talk to Leon Fuerth (LF) about his time in the White House, the role of models in policy, and intelligence analysis from a consumer's perspective. LF served as the national security adviser to Vice President Al Gore, and served on the Principals' Committee of the National Security Council, alongside the Secretary of State, the Secretary of Defense, and the President's own national security adviser while in that position. Prior to that, he worked for Gore in the Congress, spent eleven years in the State Department's Foreign Service, and served in other positions within the government over his more than thirty years in public service. He is the founder and the director of the Project on Forward Engagement, which seeks to address long-term, complex national security problems through the development of adaptive and anticipatory governance. In addition, LF is a Research Professor of International Affairs at the George Washington University's Institute for Global and International Studies, and a Distinguished Research Fellow at the National Defense University's Center for Technology and National Security Policy.

Discussion: LF's insights were quite refreshing for reasons that will be noted below. Importantly, a unifying theme bridged all of his responses to my questions, which largely weaved together a sense that the world is, and will always be, an uncertain place

where free-will exists and frustrates policy-makers' plans. What follows is a synopsis of our conversation, where answers have been combined and reordered for narrative clarity, rather than a transcript of our discussion.

The Roles of Models in Policymaking

LF's view of models and modeling was generally supportive, but different from many government sponsors or users that see models as opportunities to predict future outcomes. He was wary of the idea that human affairs could be predicted with high degrees of accuracy, and thought that efforts were better spent employing models to illuminate the kinds of unintended consequences or adaptive changes that might occur in a complex system given a particular policy choice. Additionally, he identified the value of working from models to discover particularly important levers or opportunities for influencing others.

Three challenges to the use and roles models were identified. The first challenge was practical—policy and modeling often move at different paces, and the time required developing and understanding a sufficiently rich and complex model that is credible with decision-makers may not be compatible with the speed of executive decision-making time-horizons. Likewise, increasingly sophisticated applications and explorations of models may generate more information than policy makers can digest, placing a heavy burden on new methods for distilling complex models and their behaviors into more familiar formats or new, highly efficient packages.

A second challenge philosophical—modelers and policy makers may become overly invested in particular models or worldviews that hinder their ability to develop

durable or robust policies. This over vesting occurs when models and their results are confused with real-world findings. This often occurs at the intersection between abstract theories or concepts, supported by logical models or deductive proofs, which are then seen as sources of policy solutions to complex problems. LF immediately noted that many economic models that have driven policy, such as the Laffer Curve or the J-Curve. Indeed, any model that produces a permanently valid forecast has likely been misunderstood by those advocating its application. This definition of a model overlapped with that of a theory or abstraction, which could quickly transition from a helpful cognitive aid to a misleading and dangerous source of policy solutions if interpreted literally and applied uncritically.

A third and related challenge follows from the second—if there is not a permanently valid and predictive model of the world, then policy makers and modelers must both proceed with caution. The notion that models will continuously improve and predict with greater accuracy, rather than plateau or even decline in predictive power as situations and behaviors change, leads to an uncomfortable argument against free will. People are free to adapt, change their minds, behave unexpectedly, and so on. No model can capture all possible actions and options that the actors may consider. Moreover, the implications that a model has identified the drivers of future events implies that policy makers are powerless to alter the course of events—in which case, their actions have no consequences.

Given these challenges, models can provide a more supportive role by serving to illuminate alternative perspectives and trajectories of the international system, and focus

more on understanding the sources of their dynamics than predicting outcomes.

Uncertainty and risk cannot be eliminated, but policy makers can be made increasingly aware of their sources. From this point of view, models may produce more fruitful and useful results by focusing on the exploration and illumination of potential unexpected consequences, and the countermoves of others in the system that limit policymaker's influence or power.

Being a Consumer of Intelligence Analysis

LF's perspective on intelligence support to policy provided a necessary counterpoint to the usual discussions of intelligence analysis that tend to reflect a producers viewpoint. An interesting and problematic fact of intelligence studies and efforts to reform the community in particular, has been the focus on producer's side of the producer/consumer relationship. While reforms have introduced new technologies, organizations, guidelines for personnel matters, etc., there has been little examination of how policy makers actually use and interact with intelligence analysts. There has been a one-sided focus that shifts blame and responsibility onto collectors and analysts, while shielding policy makers from not acting whenever intelligence is not deemed actionable. LF's perspective on intelligence from a consumer's perspective provided an alternative revealed how intelligence can better serve the needs of their customers.

To start with, LF noted that his career in international politics preceded his time in the White House, and included time spent in the State Department's Bureau of Intelligence and Research. He believed that this background gave him a considerable understanding how the intelligence community operated and allowed him to have a better

working relationship with the intelligence community than other policy makers who have spent their formative years in other professions and have never dealt with the intelligence community before serving in the White House or other senior policy-making organizations.

What largely separated LF from other policy makers was his placement of the burden on policy makers for the success or failure of intelligence analysis. He emphasized that the quality and utility of analysis he received was directly correlated with the quality of the questions he asked. Moreover, he felt that once policy makers became angry upon the delivery of bad news or analysis that disagreed with their beliefs the relationship could be irreparably harmed. Instead, continued dialogs and subsequent questions were necessary to provide analysts with a context for understanding policy-makers' interests, and revising collection and analysis in a way that could improve its timeliness and relevance.

This view was largely distinctive from other policy makers in that LF viewed interacting with the intelligence community as an opportunity, rather than a cost, and he noted that analysts tend to define their identity around the ability to provide policy makers with useful, relevant assessments on difficult, complex problems. Making analysts aware of the key challenges facing policy makers, allowed them to tailor their responses to his needs.

Another distinctive feature of his response related to the modeling questions as well, and that was the belief that intelligence serves to reduce uncertainty. This argument is often advanced in the intelligence studies literature, but may in fact be misleading.

Indeed, LF expressed doubts as to extent the world could be rendered knowable or predictable, so there would always remain a cloud of risk and uncertainty. The existence of such epistemological limits meant that intelligence should not necessarily provide products focused on prediction, but rather be capable of remaining continuously engaged in the policy-making process, tracking changes in the international system and helping policy makers adjust their perspectives in fluid, dynamic situations. In a manner consistent with his views on asking questions, the character of producer/consumer relations is defined by continual interaction and personal relationships, not the delivery of products.

A final point that he made was on the speed of policy making and the ability of the intelligence community to keep pace with the needs of decision-makers. Again, reiterating the importance of personal and continued relations, LF noted that the multiple layers of management in the intelligence community could often work against the interests of consumers by slowing analytic production down to the point that ceased to be timely. For example, an official response to a question could take several days to get through the entire bureaucracy, while an unofficial response may be returned in hours. Understandably, these layers ensure that the community speaks with one voice, and that analytic products bring the widest possible set of information and expertise to bear on a problem, but LF suggested that skilled analysts that were fully engaged on a problem were consistently aware of what their peers knew and thought, and appropriately and responsibly caveated their unofficial assessments. In such cases, the layers of bureaucracy slowed down the delivery of analysis to consumers and largely served the

interests of the producing organization, creating a track record and review that protected senior managers from criticism at the expense of timeliness. Other discussions with retired intelligence professionals indicate the direct relationship between analysts and policy makers is one of the most difficult and complex relationships to manage.

Importantly, his characterization of the consumer's perspective, however unconventional, reflected the conclusion commonly reached by many intelligence scholars—that the relationships between producers and consumers may be the single most important factor in determining the quality and utility of intelligence analysis and that efforts to improve tradecraft, technology, organizational structure, etc. should be viewed through the lens of this relationship.

Interview with Paul Pillar, February 1, 2012

Background: I met with Professor Paul Pillar (PP) of Georgetown University on February 1, 2012 to discuss intelligence analysis, analytic methodology, and producer/consumer relations as part of my ongoing dissertation research. The conversation was illuminating in several ways, particularly with respect to relations between analysts and policymakers. PP joined Georgetown University after a 28-year career in the US intelligence community, and also maintains an excellent blog on current foreign policy and national security issues.

Discussion: I started out our conversation with several general questions regarding the relationship between analysts and policymakers, also referred to as producer/consumer relations in the intelligence literature. I was particularly interested in the ways relations between the two affect how analysts go about their work, and select methodologies in their effort to address the needs and interests of policymakers.

PP noted much of the challenge for intelligence analysts is a result of the busy and dynamic nature of policymakers' responsibilities. While many of the idealizations of intelligence work assume that policymakers can articulate their intelligence needs to analysts and collectors, he noted that reality is far more complex. Policymakers' ideas, interests, and thoughts about issues can evolve quickly and unexpectedly. This makes it difficult for them to anticipate precisely what their intelligence needs are or will be in the future. Moreover, because policymakers have little time, they cannot always educate themselves on the issues that intelligence analysts have expertise on. Alternatively,

intelligence analysts cannot replicate the dynamic demands on policymakers, which are critical to understanding how consumers' ideas, needs, and perspectives will evolve.

A follow-on question was on the differences between what qualifies a high-quality finished intelligence product within the intelligence community compared amongst their policymaking consumers. PP noted that internally, features such as the intellectual cogency and elegance of an assessment are deemed important. Alternatively, consumers may care more about the usefulness of the assessment, and wonder why it could not have been produced sooner, or if anything could have been done to resolve outstanding uncertainties and gaps in information.

One of PP's most insightful comments came as he emphasized the primacy of consumer interests. Importantly, he noted that tradecraft and methodology increased in importance whenever analysts received less guidance from policymakers about their interests and requirements. Thus, analysis is often a straightforward act when sufficient guidance regarding priorities and interests are provided. However, in the absence of these things, analysis becomes increasingly subject to methodological considerations in order to expose the assumptions and logic behind inferences given their uncertain relevance to consumer's needs. PP's note on the importance of methodology as a substitute for external guidance also led to a related but inverted observation—that increased customer interest and guidance in a subject often reduced the extent to which analysts could explore alternatives.

PP's observations suggest an interesting tension that is not unfamiliar to evolutionary and complex systems that must balance the use of resources between

exploration and exploitation. In this case, too much interest and guidance from consumers focuses analysts on exploiting a limited set of information and perspectives that will satisfy the interests of policymakers at the expense of searching for alternative points of views and additional collection. As a result, analysts may lose the ability to explore alternatives (spending resources on potentially uninteresting or unproductive lines of reasoning), missing out on potentially important discoveries or opportunities to reframe questions. By comparison, if policymakers take no interest in an intelligence problem, then analysts may constantly explore alternative frameworks, gather data, and engage in an endless search process that may produce conflicting assessments that do not build toward a shared understanding of the issues or character of a target or opportunity. Thus, too little exploration leaves analysts and policymakers vulnerable to surprise, while too much exploration does not provide any consistent theme or approach around which an analytic foundation can be developed that will warrant the sustained interest of consumers.

A related question that I asked PP had to do with the contributions of Structured Analytic Techniques (SATs) (formerly known as Alternative Analysis). SATs have been introduced into tradecraft in an effort to make analytic assessments more transparent with respect to the assumptions of the analysts, the data they used, and the logic of their arguments—making the relationships between evidence and inference explicit. Additionally, they seek to extend the number of perspectives under consideration, ensuring that analysts reduce the psychological tendency to satisfice and accept the first theory that explains a particular situation or answers a question satisfactorily. On the

whole, PP believed that such techniques were generally helpful and should be encouraged, but also observed that their implementation was not simple or straightforward. Specifically, it was noted that particular pieces intended to stimulate thinking or advance a case for a given hypothesis, regardless of its merits, could invite cherry-picking. PP noted examples of CIA analysis during Vietnam, where a particular assessment was provided to the White House to answer a question regarding whether a particular case for success could be made, and was then treated as a considered analytic judgment once stripped of the caveats that characterized it as an intellectual exercise. Likewise, he referred to a similar piece was written regarding an effort to assassinate Pope John Paul, laying out how to interpret the available intelligence to make a case for Soviet covert action—yet, PP noted that no consumer had asked for an analysis setting out the case against Soviet sponsorship of the assassination attempt. Thus, it was suggested that in some cases, SATs and extensive explorations of alternatives could exacerbate certain producer/consumer pathologies, and that when such explorations were beneficial, packing them in a larger, comprehensive product could hedge against cherry picking single perspectives or scenarios if presented as individual products.

PP argued that the successful employment of SATs required managerial engagement and support. Because of the agency's effort to advance a corporate analytic line, managers needed to be disciplined regarding how and when alternatives were presented to consumers and ensure that Devil's Advocacy, Red Cell, Scenarios, and other products were developed and advanced in the context of corporate perspectives and needs. Absent managerial support, analysts could not advance products without risking

isolation or employing resources in unproductive ways. This suggested that a bottom-up approach to analytic production was unlikely to improve the overall performance of the community. Finally, PP noted that there remain limits to what tradecraft and methodology can offer, and that achieving the levels of predictive success that many outsiders desire may not be possible.

On the issue of Analytic Integrity Standards and PDB guidance, PP did not feel that internal editorial practices significantly affected analytic work and were generally positive. Specifically, they helped to ensure that the policy preferences of the analysts themselves did not shade their judgments, and made sure that the use of language employed by analysts was consistent across products—guaranteeing that everyone knew the “strike zone.”

In regards to the relationships between collectors and analysts, PP noted that this was an area where the community has significantly improved over the last couple of decades and that communications and increased flows of information across boundaries had helped a lot. He noted that analysts have increasingly supported collection through the use of targeting analysis, and that collectors have played a stronger role in analysis by validating the use of data by analysts, ensuring that its use is appropriate based on the context of its collection (PP specifically noted the case of Curveball as a case where analysts would have benefitted from the guidance of collectors in order to understand how to use/not use data). On this topic, he noted that the relationship is generally good, but unstructured and informal and that the use of computational simulations might add

structure and sophistication to this relationship and reduce the use of scarce resources by limiting the amount of unnecessary information pulled into the system.

The final point about unnecessary information served as a prelude to questions about technology and its role in analysis. PP argued that the most important challenge facing analysts was the massive volume of information that must be processed and examined, much of which coming from open sources. He argued that this problem would become increasingly acute as human resources become scarcer and that simply providing analysts with the tools they want would not necessarily solve the institutional problems created by the massive quantities of data that required evaluation.

Another question that I asked dealt with the question how intelligence analysts' work distinguishes them from scholars who study similar topics. PP noted that analysts' closeness to policymakers had a significant impact on their work processes, and resulted in elevated expectations, the need to address policy relevant concerns in real-time, and professional and personal pressure and responsibility for failures. He noted that no scholar risks the headline "scholarly failure" if their predictions fail to materialize.

Interestingly, PP's emphasis on constant contact and need to support consumers evoked another evolutionary model advanced by Herbert Simon and his work on complex systems. Simon used the analogy of two watchmakers, each constructing watches with the same number of parts. The difference in their work, however, was that one assembled the watches as a single unit, and, therefore, if disturbed, would lose his work and have to start from the beginning. By comparison, the second watchmaker developed watches in modules, such at any disruption would only affect the module under development,

leaving those that were already constructed unaffected. The result was that the second, modular watchmaker was more productive.

This metaphor may capture some of the salient differences between intelligence analysts and scholars. Because analysts are constantly in contact with policymakers, they are subjected to many “disruptions” and must therefore do their work in smaller pieces and cannot wait until the “end” in order to reach a conclusion or advance a judgment. This may also mean that analysts’ views or beliefs may appear to be inconsistent over time, simply because the evolution of their thinking will be transparent and displayed because of the need to constantly publish their best estimate at that moment in order to support consumers. By contrast, scholars may hide their evolutionary path from others, meaning that what appears as a single, integrated coherent and consistent final result may be the product of their internal learning and deliberations that remain private or shared only among peers, but never subjected to demands or disruptions of outsiders during the research process.

A final question concerned the prospects for new models of producer/consumer relations. Specifically, new models have been advanced such as consulting, net assessment, or even joint production and assessment. PP felt that the limitations of new models were already known and that preserving the distance between producers and consumers remained an important and necessary part of intelligence in the future. He pointed to the conclusions of the Butler Report that suggested the two were too integrated in the UK in the buildup to the Iraq War. There was an acknowledgement that wargames and simulation might serve as a bridge between the two and allow each to work together

without going as far as “joint publication” which would necessarily politicize intelligence. PP also noted that policymakers who had backgrounds in intelligence were different consumers, more informed about intelligence capabilities and processes, but were very few in number.

Interview with John Hanley, February 9 2012

Background: Discussion with John Hanley (JH), Director of Strategy for Office of the Director of National Intelligence. Prior to joining ODNI, JH served as an officer in the US Navy and held senior positions in the DOD. I asked John broad questions that primarily focused on three topical areas – the difference between analysis and analytic communities within the DOD and Intelligence Community (IC) based on his experiences, his perspectives on the current state of the IC given his role and position within the ODNI, and general comparisons between intelligence analysis and academic scholarship. Note: since the time of this interview, Dr. Hanley has retired from the ODNI.

Discussion: I initially asked JH about the differences and similarities between analysis as performed within the DOD/military and the IC given his experiences in both environments. He noted that his background as a submariner was instructive within the development of DOD's analytic capabilities and the history of Operations Research (OR), which constituted an effort to apply scientific methods (and practicing scientists) to the problems of Anti-Submarine Warfare (ASW). He noted that the ASW community went from having little analytic capability to becoming one of the premiere users of OR methods after a dedicated, and sustained organizational commitment to the problems posed in WWII and the Cold War.

When compared with the IC, particularly the CIA, there was little comparable story with respect to OR. JH attributed this to differences in analytic cultures and noted that the DOD (particularly the analytic and managerial organizations that were legacies of McNamara) was largely a numeric culture while the IC was predominately as literature

one. Thus, investments in particular analytic tools and methods that aligned with quantitative and formal modeling, which often found fertile ground in the DOD, were not necessarily long-lived within the IC. JH was quick to note that this was necessarily because IC analysts lacked the ability to work with such approaches, and indeed, he noted that analysts often moved between communities, but that the organizations themselves emphasized different questions and served different customers. Nevertheless, JH believed that the IC could benefit from using more formal analytic methods when problems warranted it. However, JH also noted that DOD had often become overly reliant on formal methods in several cases, particularly large combat simulations, and could often be overconfident in the models and prone to being misled by modeling results if not sufficiently familiar with their inner workings and embedded assumptions.

I asked whether the differences between the DOD and IC affected the roles and responsibilities of managers themselves. JH felt that from an organizational perspective that managers had very similar jobs regardless of the organization, but that differences often existed based on their experiences. Thus, differences could be seen, but they were the result of personal experiences rather than organizational demands.

Because it was clear that a great imbalance existed regarding the use of formal models between the DOD and the IC, I wanted to know if this applied to the larger class of simulations, which also included wargames. JH felt that as exploratory tools, games served a similar role in both communities, but that the DOD had several distinct advantages when dealing with military issues. Specifically, because military questions often involved the development and use of weapons systems, many of the questions

raised by a game could subsequently be explored through exercises and concerted data gathering efforts to validate modeling results against real-world outcomes. Likewise, questions about parameters concerning operational or technical variables could also be determined empirically. Thus, a tighter loop and feedback mechanisms could support wargames on the DOD when compared with similar games in the IC. So, even though wargames played similar roles with respect to analysis within the DOD and IC, the character of the games, particularly with respect to their fidelity or ability to be linked with operations or supported by formal models, was greater on the DOD side as a result of combination of culture and resources.

JH's characterization of the linkages between analytic or exploratory wargames and models with operators prompted me ask about the overall relationship between research and development, analysis, and operations. JH again drew upon his navy experience and discussed how important close relationships between the three really were and that many of the major problems affecting the DOD today were a result of this relationship breaking down and becoming overly rigid and tied to requirements processes that inhibit groups from working together. Specifically, he discussed the development of the AEGIS missile defense system, where R&D and operators worked closely together in an evolutionary process that explored how combinations of technology and operational use could push radar performance past previously anticipated performance limitations, resulting in a system that would have never been envisioned by the requirements process. JH noted that this process was heavy vested in tight feedback loops, the collection of empirical data on operational performance, exploration of the possible (as opposed to the

emphasis on the feasible that results from the requirements process), and a minimized role for contractors with the majority of the design, engineering, and testing performed in-house.

JH's discussion of the AEGIS system denoted the importance of evolutionary approaches to problem solving where prototypes are developed rapidly, and then experimented with in operational contexts. He noted that such an approach was increasingly scarce given the combination of DOD bureaucracy and Congressional oversight however. In such cases, the IC and the classified Special Access Programs (SAPs) that operate with greater secrecy, and therefore less transparency, provide greater opportunities for experimentation and risk taking in accordance with the evolutionary model.

After discussing some of the features of technical development, I focused on questions that were more specific to the IC and the ODNI itself. My first question was about how the IC evaluates the quality of its analysis and analytic products. JH noted that it was a very difficult problem and something that was still in its infancy with respect to developing evaluative criteria beyond common production metrics. He emphasized the recent development of analytic integrity standards, and a greater emphasis on comparing prior analytic forecasts with actual outcomes systematically. JH felt that greater attention needed to be given to methods and analytic approaches, specifically focusing on performing cost-benefit analyses in order to determine whether different techniques are better than others with respect to generating useful insights for consumers. I asked JH about the dangers of such assessments, particularly analytic integrity standards, and he

noted that there is always a risk that well-intentioned efforts to guide producers may become check-lists that are employed without consideration of context. JH also noted that the analytic integrity standards could be seen as attempting to cope with the challenges of discerning arguments based on evidence vs. those based on intuition, and that the standards favored or promoted analysis grounded in evidence over those based on logical argumentation. From the perspective of modeling and simulation in the analytic process, such as distinction can be problematic, given that models are products of intuition/theory/simplifications, yet generate data upon which inferences are based, and thus grounded in empirical observations of artificial systems.

Another question I asked JH related to new models of intelligence support to policymakers, such as increasingly considering blue capabilities and actions in assessments of foreign actors, the use of wargames or simulations with policymakers as players, etc. JH noted that there a shift has occurred in the policy world and that expectations are increasingly shaped by social media technologies that position the user/consumer as an interactive participant in the delivery of information. He noted that providers of information are increasingly faced with very rich and fast feedback from users that expect to have influence over the production process—shaping its format and content. As a result, the prospects for using simulations as an analytic tool that allows for intelligence and policymakers to collectively explore complex problems may be rising in importance and effectiveness as a means of producing and delivering intelligence analysis (or perhaps intelligence “experiences”?).

The use of simulations as providers of experience that exercise policymakers and their staffs was also mentioned by Leon Fuerth in another discussion. Leon was concerned with the time required to develop models, simulations, and games that were suitable to the needs policymakers. When faced with the same question, JH acknowledged the concern, but was more optimistic, noting that game and software developers were increasingly employing specialized tools for developing products that significantly cut down development and testing time (indeed, many video games have now provided players with content creation tools for developing their own modifications and content, and these tools require no programming skills). Thus, JH believes that the infrastructure to support modeling and simulation is becoming increasingly streamlined and less complex, and these innovations will reduce the amount of time required to develop models and simulations for use in analysis or with policymakers.

However, JH also observed that the abstract nature of models will always impose some limitations on the extent they can be expected to replicate reality. As a result, games and simulations will be less helpful in crisis situations when policymakers' concerns are focused on very specific courses of actions and options. Instead, as questions are more strategic and focused on long-term concerns, models and simulations can play a more influential role by helping to define the characteristics or properties of sound strategies, identifying vulnerabilities, or conditions for opportunistic action. On this issue, JH used the phrase "striking while the iron's cold" to highlight the importance of interacting with policymakers when issues are not urgent or pressing and rational and critical analysis can be influential. Thus, on many strategic issues, analysis must be

performed in advance and may even need to sit dormant until the interest in the question/problem is raised.

JH concluded that many of the advancements in tradecraft were being made but that innovations were largely confined to small groups and not encouraged systematically. He repeated his belief that greater attention to assessing tradecraft and evaluating techniques was needed in order to give managers and analysts a better understanding of what works, what doesn't, and how to employ their resources in the production process.

My final questions were about the similarities and differences between intelligence analysis and scholarship. JH noted that academics tend to be more analytic than synthetic – focusing on the isolation of particular variables or causes in order to examine and explain events rather than aggregating them to obtain a comprehensive understanding. The result is that much of scholarship, even from policy-oriented academics, has been unhelpful to real-world policy and analysts because of their emphasis on studying the “lens” or theory and its applicability to many problems or cases, rather than the particulars of specific problems. This again, was reminiscent of John Lewis Gaddis’s argument about “particular generalization” in which argues that lessons should be drawn from the study of individual cases and examples, rather than envisioning a theory and then attempted to fit multiple cases to it.

Interview with Carmen Medina, June 5 and July 3, 2012

Background: Carmen Media (CM) spent more than 30 years at the CIA, and served in multiple leadership positions—the most notable of which included serving as the Deputy Director of Intelligence and Director for the Center for the Study of Intelligence. After retiring from the CIA she joined Deloitte, where she continues to support the Intelligence Community (IC) and other business activities including Deloitte’s Center for Federal Innovation. My interview with CM spanned two sessions that included an in-person discussion followed up with a series of questions addressed via email correspondence. What follows is short summary of our dialog that primarily summarizes the focused email correspondence.

Discussion: My first question to CM concerned the ways in which intelligence analysis has changed as a discipline over her time in the community. She argued that a fundamental change had taken place in which analysts previously had a paucity of data to work from in their efforts to understand different situations but now drown in data, particularly the “digital exhaust” left behind by actors embedded in the modern, global information infrastructure. CM noted that as a result, questions and answers have also changed from focusing on uncovering the secrets and intentions of closed societies towards discovering essential “truths” or characteristics about individuals and large groups, many of which they may be unaware of themselves.

CM also noted that along with the changes in the quality of the data that is available to intelligence analysts, the quality of that data is changing too. Thus, she argued that many of theories about social behavior and dynamics are being challenged by

the kinds of data that are now available, e.g. detailed microlevel, longitudinal records on individuals and the distributional properties of groups (thus looking beyond the mean/median) which allows for a reexamination of old theories as well as a demand for new ones that can better align with observed reality.

My next question asked about the character of intelligence questions themselves and how they have changed. CM responded that the focus on nation-states has become tiresome and often misleading or at least too slow to pick up on trends that could be seen sooner and more clearly from different perspectives. Indeed, she noted that so many of the questions analysts and policy makers now confront rarely deal with the activities or capabilities of nation-states in isolation of other factors, whether terrorist groups, the environment, regional economies, etc.

I also asked CM about the role of technology in intelligence analysis, both from the perspective of performing analysis and also delivering products to consumers. She noted that the IC's reluctance to embrace emerging technologies has left them unable to fully appreciate the ways in which technology is changing society itself, e.g. the ways groups organize and task themselves to accomplish their goals. She also noted that she hated the way the community thought about their work as producing and delivering "products" to consumers. She argued that most are unnecessary and rarely viewed or used by policy makers. Instead, she argued that IC should have two streams of analytic activities: 1) a forum where consumers can ask questions of analysts and receive answers tailored to their needs, and 2) a standing deep research activity that generates new

insights. She noted that research efforts may take years to perform, yet be summarized in a single short paragraph.

Given CM's career, much of which was spent in senior managerial positions, I asked how her perspective changed as she transitioned from being an analyst to a manager. CM responded that her core beliefs about intelligence analysis did not change much over time, although she became increasingly sensitive to the extent to which traditional definitions of national security had to be rethought. She also noted that as a manager she came to see her role as one of overseeing the government's sensemaking network, and that she needed to think about the performance of the network as a whole, rather than focus on the performance of individual analysts.

I also asked CM about the uses of simulations within the IC (both manual wargames and formal models), and the ways in which consumers' expectations of intelligence affect the selection of analytic methods. She noted that wargames were the most prominent type of simulation used in the community, but that their contributions were rather indirect. CM argued that the games themselves were often too subjective to drive analysis, but they served a valuable social role by encouraging participants to have the conversations they needed to have anyway.

CM also noted that consumer's questions and expectations do affect analytic methods. She argued that too many policy makers expected answers from analysts and wanted the intelligence community to "make the call." However, she noted that these demands were fundamentally unhelpful, and that policy makers should expect the intelligence process to help them become clearer about issues rather than discover truth.

She believed, however, that the harder policy makers pushed for concrete answers, rather than engage in a sustained process of discovery and learning, the more they encouraged dishonest behavior by analysts seeking to satisfy the demands of their consumers.

My final question for CM asked about her most challenging and rewarding experiences of her time in the IC. She noted that the greatest rewards for her came from helping others become better thinkers and discovering and exploring new ideas.

Interview with James Bruce, August 13, 2012

Background: James Bruce (JB) spent 24 years in the intelligence community, serving in senior positions in the National Intelligence Council, the Silberman-Robb WMD Commission, and the CIA's Sherman Kent School. After retiring in from the CIA in 2005, he joined RAND as a Senior Political Scientist. He is also an adjunct professor at Georgetown University's Security Studies Program, and is coeditor, with Roger George, of *Analyzing Intelligence: Origins, Obstacles, and Innovations* published by Georgetown University Press.

Discussion: On August 13, 2012 I met with JB to discuss intelligence analysis, the use of modeling and simulation, and many of the underlying challenges of analytic tradecraft and production from the perspective of the philosophy of science. Afterwards, our dialogue continued via email, introducing new topics and refining points made previously. Because of the wide ranging character of our continuing discussion, and the fact that it is a two-way conversation rather than an interview, the summary below touches upon four of the most interesting topics that we discussed rather than a detailed reconstruction of entire discourse.

The first topic that we discussed concerned the characteristics of useful intelligence analysis and whether the practice of intelligence analysis could be consistent with science. JB argued that useful analytic products offer consumers judgments, forecasts and insights, each of which constitutes a different type of information. For example, the intellectual justification for reaching a judgment about the capabilities or intentions of a foreign missile program may be quite different than forecasts regarding how scientific discoveries and engineering applications might produce strategic

consequences that alter the balance of power in the international system. During our conversation, the development of forecasts and insight received the bulk of our attention due to the difficulties associated with each when compared with providing consumers with judgments. Indeed, judgments were relatively straight forward to identify and produce, although determining whether they are based on sound evidence, logical reasoning, and are robust or fragile in light of uncertainties are a continuous problem.

JB argued that at the foundation of intelligence production were two major scientific challenges—the generation of hypotheses and their testing. His treatment of intelligence as a scientific practice was consequential in two ways. First, he challenged the treatment of intelligence as a strictly artistic or intuitive act, by holding analysts to the procedural standards of the scientific method as characterized by scientific positivism. Second, JB's inclusion of hypothesis generation into the scientific method marks an important difference between his thinking and the work of others that have advocated scientific approaches to intelligence analysis, e.g. Ben-Israel who argued that where theories came from was outside of the boundaries of science or scientific consideration. By including how hypotheses are generated, JB placed the act of theorization itself into the scope of what can be treated and studied systematically, and exposed the development of application of analysts' mindsets or mental models to evaluation with respect to rigor and logical coherence. This is of great importance from the perspective of modeling and simulation, because it introduced opportunities to employ machine learning, evolutionary computation and other approaches that focus on the discovery of new models and relations between variables and actors, rather than limit the role of models to hypothesis

testing alone. Indeed, even without the heavy use of computational tools, the very act of model development and design often serve an important role in theory development and the generation of hypotheses that guide future testing in scientific or analytic practice.

Our second topic flowed logically from the first. By treating analysis as a type of scientific hypothesis testing, JB noted that several of the long-standing challenges in the philosophy of science emerge and affect intelligence analysis. Specifically, he noted the important differences between nomothetic and ideographic approaches to science, history and analysis. JB noted that in the former, empirical cases are taken to be instances of a general phenomenon that can be studied and tested. Alternatively, the latter sees history as a single path through a dynamic system, where each case is unique and embedded in a context determined by prior experiences and perceptions of earlier events. In the ideographic case, the empirical record is no longer a collection of independent events that provide the basis for general claims, but a particular trajectory that may or may not be indicative of repeatable outcomes, e.g. if the Cold War between the US and Soviet Union were replayed 1,000 times, how many different ways would have it concluded and with what frequencies? Our conversation noted that each approach has important implications for what it means to test hypotheses and how to test them. The nomothetic approach is consistent with the dominant approaches found in the social sciences, e.g. Large-N statistical and Small-N case study approaches, while the idiosyncratic approach is more consistent with history, ecology, evolutionary biology, and other complex systems analysis. We discussed how simulation might provide opportunities to bridge gaps between the two approaches, giving analysts and consumers

greater confidence in assessments. I noted that Agent-Based Modeling, specifically because of its ability to capture and represent microlevel differences between cases, may provide a powerful tool for differentiating between those outcomes that are structurally determined or based on variables that are recurrent across cases (nomothetic), and those outcomes that are contingent and dependent on the specific or unique features that make one situation unique from others (idiosyncratic).

A third topic of discussion was on the characteristics of knowledge and foreknowledge about the world. JB noted that intelligence is often defined as foreknowledge about the world. He started with a definition of knowledge that he found particularly helpful in his own thinking, where knowledge was defined as justified, true belief. We discussed the implications of this definition, particularly as it applied to intelligence given the treatment of intelligence as foreknowledge about the world and therefore constrained by what is knowable *a priori*. Importantly, the boundaries between what is epistemologically knowable or not is contingent upon whether one views the world in nomothetic terms or idiosyncratic ones because each implies a different belief about the repeatability of events and the reliability of patterns across cases.

Additionally, we discussed how in many cases a notable gap existed between what could be known beforehand as a practical matter and necessary role of decisions and actions beyond those of the intelligence community in determining the truth about assessments. For example, the foreknowledge of Osama Bin Laden's (OBL's) location in Pakistan was justified based on intelligence observations of the Abbottabad compound and believed by analysts and policy makers in some probabilistic sense (each person who

knew about the compound may have believed the identity of its mysterious resident was OBL with a different level of confidence), but the truth regarding his location could only be determined by physically raiding his compound and determining the identity of the occupants. Thus, the intelligence community may not always be capable of determining the truth of its assessments and may require policy makers to commit resources and authorities to operations in order to ultimately validate the truthfulness of intelligence assessments.

Our fourth topic gave additional consideration to the problem of hypothesis testing in intelligence analysis. JB noted that the means for hypothesis testing when dealing with quantitative questions are well established and often quite accurate. Thus, on quantitative matters analysts may have sufficient opportunities to develop knowledge (again defined as justified true belief) based on the maturity of scientific hypothesis testing particularly through the use of statistical inference. However, JB noted that the situation was far less developed for testing qualitative hypotheses, which constitute the bulk of intelligence questions. Thus, the strength of justifications provided for many beliefs about intelligence targets may not be knowable given the inability to adequately test qualitative hypotheses. JB noted that qualitative assessments continue to pose three important challenges.

First, precise characterizations of uncertainty regarding the qualitative assessments remain problematic. While analytic tradecraft emphasizes providing consumers with critical uncertainties that would affect analytical judgments, precisely how to do so regarding qualitative matters remains an unsolved scientific problem.

Second, JB noted that in many ways, intelligence analysis can be reduced to a series of logical sentences that are either true or false. This approach would seem to owe its origins to logic, philosophy and linguistics, and I do not know if intelligence analysis has been evaluated in this fashion before. I suspect that the closest approaches to this sort of logical construction of intelligence judgments would be the employment of Bayesian techniques to construct assessments out of chains of conditional probabilities. Finally, JB noted that mapping the space between facts and judgments remained a considerable philosophical problem. He argued that the structure of this space was important for determining where knowledge resides and how it is constructed out of the many source and methods available to analysts?

Interview with Jennifer Sims, August 20, 2012

Background: Jennifer Sims (JS) is currently working on a grant to continue to develop her theory of adaptive realism, which incorporates intelligence activities into international relations theory. She was previously the Director of Intelligence Studies at Georgetown University's Security Studies Program. JS also served as Deputy Assistant Secretary of State for Intelligence Coordination, the Department of State's first Coordinator for Intelligence Resources and Planning, and on the staff of the Senate Select Committee on Intelligence.

Discussion: On August 20th, 2012 I interviewed JS to discuss her theory of adaptive realism and the intelligence community more generally. As with other interviews, our conversation was far reaching, open and discursive. Thus, the following summary presents several key points that came up in conversation rather than provide an exhaustive transcript of our exchange.

One of the first questions I asked JS was about producer/consumer relations. She noted that the community must adopt a wider perspective regarding its role in policy making and that without the authority and expertise to examine US actions, goals and capabilities it cannot adequately determine what developments in the international system are truly threatening or identify opportunities to be exploited. Thus, JS argued that the intelligence community must increasingly perform net assessments and broaden their analytic focus to include assessments of the relative strengths and weakness of the US (blue) and foreign actors (red). From her perspective, the context of how the US and others compete or interact is crucial for fulfilling the analytic mission of the community,

and this context cannot be established without carefully studying and considering blue as well as red.

Her emphasis on interactive context is important from a methodological perspective, particularly when considering the employment of Agent-Based Models (ABMs). Whereas the traditional analytic paradigm within the intelligence community is to study capabilities and intentions, ABMs place actors into an interactive and often strategic context that has its own independent effects on the system. Factors such as the order in which agents act can affect outcomes, despite the fact that they are not properties of agents' capabilities or intentions. Indeed, in many models, agents with homogeneous intentions and capabilities can nevertheless produce ranges of distinct outcomes, indicating that other features of the system are in play and shaping results. Thus, ABMs might play an important role in broadening intelligence analysis and collection by adding the interactive structure through which interactions occur to the list of actor specific capabilities and intentions already examined.

JS's emphasis on net assessment also revealed the need to expand the reach of the intelligence community to include nontraditional, domestic consumers. She noted that one of the problems with strategic warning prior to 9/11 was that the National Security Advisor, Condoleezza Rice, had no domestic constituency and therefore was unable to alert parties outside of the foreign policy community, such as the FAA, about the threat posed by al Qaeda. JS noted that a richer set of intelligence community consumers, especially those now embedded in Department of Homeland Security, would have given

intelligence analysts and policy makers greater flexibility to respond to the threat posed by al Qaeda.

JS's view of intelligence is distinctive due to her emphasis on strategic interaction, secrecy, and deception. Both she and Jim Bruce (interviewed the week earlier) have given considerable attention to these topics, particularly secrecy, denial and deception, and the integration of intelligence collection and analysis, yet each has emphasized different aspects of analytic efforts. Whereas Bruce focused on epistemology, the scientific method and hypothesis testing in order to understand the quality of analytic judgments, forecasts, and insights, JS has focused on whether intelligence information provides decision makers with a competitive advantage relative to their rivals. The implications of their differing emphasis is stark—Bruce's concerns are about how we know whether intelligence estimates are justified based on empirically verifiable truths, while JS's concern is whether intelligence consumers possess a relative advantage over their rivals, accepting that both may possess inaccurate or illusory perspectives.

As we spoke, JS and I repeatedly returned to the subject of strategic intelligence, particularly its importance to consumers and difficulty to produce. She noted that many changes were required within the intelligence community in order to redress existing deficiencies.

First, JS noted that analytic methods should first and foremost concentrate on the relationships between producers and consumers. This means placing matters such as achieving the most accurate assessment secondary, behind building trust with consumers

and getting them to accept analytic inputs. Importantly, this does not mean that matters such as truth or accuracy are unimportant, but rather that the development and use of strategic intelligence is a dynamic process whose sequence first requires producers to earn the trust of consumers before they will accept their inputs and give alternative, challenging perspectives consideration.

From JS's perspective, the objective of strategic intelligence is to get policy makers out of their existing paradigms and consider alternative perspectives on world events and potential futures. This requires the intelligence community to provide mechanisms for preparing consumers to think through the implications of potential events and warn of their prospective occurrence. This is achieved through the integration of collection and analysis, where each informs and guides the other. Thus, collection tests analytic hypotheses by searching for indicators that would confirm or refute expectations, while providing the information needed to specify, parameterize or calibrate analytic models (both formal and informal) in order to ensure that analysts work from the best available representations of reality. Alternatively, analysis is needed to specify and prioritize intelligence targets against which to focus collection and provide interpretive frameworks for making sense of the data that collectors provide. While strategic intelligence is often considered an expansive act that proliferates the number of scenarios and perspectives to present to consumers, JS noted that the effective integration of collection and operations can also limit the extent to which the intelligence community and policy makers must consider alternative perspectives or worldviews by continuously ruling out developments for which no combination of theory and data can be found. Thus, at the nexus of analysis

and collection, intelligence producers continuously develop and suggest alternatives for consumers' consideration while simultaneously eliminating previously identified possibilities for which there is no evidence to merit continued attention.

A second strategic intelligence issue identified by JS was the movement from estimative to anticipatory intelligence. This largely means transitioning from the delivery of forward-looking yet static assessments of foreign activities and developments, towards a dynamic and interactive evaluation that includes consumers gaming how they might respond if a particular scenario comes to pass. JS argued that successful anticipation is based on starting from speculation to derive the implications of potential scenarios and associated indicators of their occurrence. Once strategic threats (or opportunities) have been identified, getting consumers to buy into analytic assessments about their significance is essential. Afterwards, the intelligence community can search for indicators of different strategic developments and warn policy makers of their occurrence. By placing consumers between the development of scenarios and the search for their occurrence, JS ensured that warning would be accepted by consumers when delivered by intelligence producers because of their prior exposure to the framework against which warning is provided.

As JS and I continued our discussion, we talked about the potential for modeling and simulation to make significant contributions to intelligence analysis. Because of her views about strategic interaction and bounded rationality, already central to her theory of adaptive realism in international relations, she understood a great deal about how game theory could be applied yet was sensitive to the cognitive biases and limits of

organizations and individuals. Because Agent-Based Modeling provided opportunities to simulate strategic interactions, including deceptive behavior and the use of private or secret information by decision makers, she acknowledged they could make a contribution, especially in cases where coevolution, learning, and adaptation were the drivers of competitive balances and advantages.

A final topic covered in our discussion concerned the application of models more generally in analytic processes. One way was to focus on gathering sufficient data to properly calibrate or tune models in order to provide projections of contemporary circumstances, or at least precisely specified initial conditions. A second way to employ models was to search broadly across the parameter space in order to identify what outcomes are correlated with what inputs, and whether there are features that are insensitive to parameter choices and structurally determined vs. results that are contingent on specific combinations of inputs or sequences of events. JS noted that both approaches had their merits, but believed that linking models to data would be essential to earn the trust of consumers.

Interview with Joseph Eash, September 6, 2012

Background: Joseph Eash, III (JE) had a distinguished career in the US Air Force and Pentagon, where he led the development of advanced technologies for many decades. He served at the National Reconnaissance Office, for which he received the Pioneer Award, served as the Deputy Undersecretary of Defense for Advanced Technology, and retired from government in 2005 after serving as a Chief Scientist for Computational Social Science Modeling at the Center for Technology and National Security Policy at National Defense University.

Discussion: I asked JE several questions about his time in government, where he served in intelligence, technology, and policy positions. As a result, he had a unique set of experiences upon to draw upon when considering the development and integration of new analytic technologies and methods into the intelligence community.

JE started by noting that we will never perfect information or intelligence, but that the ranges of what can happen can be understood by understanding technical things and the processes that people follow to develop artifacts or execute plans. What we do is spend time developing templates or patterns of how activities are performed, and then sample from the world in an effort to see if we can fill in the templates.

JE noted that his characterization of intelligence introduced several challenges. First, the problem of templates was largely a theoretical one. Templates are essentially characterizations of how something might be accomplished, whether one seeks to develop a nuclear weapon or run a political campaign. In a sense, they might be considered algorithms that constitute a set of steps to be followed in order to achieve a

particular goal. Collection then samples from the world and attempts to match observations against templates in order to understand what is happening and what might happen. This means that intelligence requires both imaginative, theoretical work to develop libraries of templates, and also an understanding of collection systems and limitations in order to precisely understand and communicate what information is gathered and what is not. Importantly, JE noted that in this context a single piece of information or intelligence has very little meaning – what matters is how it relates with other information in order to fill in a template.

JE noted that in many cases we look for things with no template, or an unknown template. This leads to problems regarding what to do with collected information. We stretch it, exaggerate it, do a bunch of things to make sense out of it, but absent a general template reflecting a robust understanding of people, processes, etc. we are likely to misunderstand what we gather.

JE continued by noting the importance of cognitive factors in the analytic process. He discussed the transition from film to digital images within the intelligence community. He noted that one of the curious aspects of analysis was how the brain affected the perception of images. When images were obscured, analysts often showed greater confidence in their ability to identify its contents than when it was clearer. Moreover, once an analyst committed to particular identification or interpretation, it was very difficult to back them off their conclusion, even as the image was made sharper, the brain got stuck and attached to a particular interpretation.

JE believed that the cognitive challenges of interpreting imagery were representative of a larger range of analytic challenges. He noted that in many cases, the available information is often of low quality, akin to the obscured image, but analysts must make judgments from them, and this process can lock errors into the analytic process. He also noted how important good, high quality data is early in the analytic process. Because analysts were more accurate when first shown a high quality image than a low quality image, but often resisted revising assessments from low quality images after being shown a high quality image, an asymmetry existed where more information was needed to change an observer's mind than to shape it in the first place.

JE argued that one of the major problems in intelligence and policy is that analysts and decision makers are making assessments and policies with mental models or templates of unknown quality, and that the information going into them is also of poor or unknown quality. As a result, everyone is vulnerable to seeing what they wish to see and unless challenged by high quality information, their perceptual biases are likely to carry the analysis. He continued by noting that in many cases, the underlying theoretical work required to develop templates has not been done, and that the information we can collect cannot justify the conclusions we are trying to reach.

JE compared the problem of analysis to stealth in the radar world. He noted that collecting against a stealthy vehicle was a difficult problem – collection systems only get glimpses of the target, and these glimpses could be quite difficult to integrate into a common picture, e.g. were they of the same vehicle or multiple vehicles? In the abstract, the problem was the same – how to get the most out of low-resolution

information? Better collection was one possibility, which meant collecting with a better sample rate or resolution. A second solution was to develop new templates that would allow analysis to proceed with less information.

I followed up by asking JE about the differences between intelligence problems in the technical domain and the development of new technologies compares with social science problems that became increasingly important during the latter years of his career. JE noted that the social domain extended the range of considerations and duration of time to be concerned with, but also the opportunities to develop an understanding of the international system. He noted that in a simple case, e.g. determining if an attack was occurring, one could look into the technical domain to look at the horizon using optical, infrared, acoustic, and other sensors. However, extending into the social domain opened the problem up to provide new indicators for warning, many of which would be sooner than the technical ones, e.g. the execution of organizational procedures for attack preparation or social indicators such as communications between family members.

JE argued that the social dimension provided new dimensions across which signals could be correlated to gain a more complete and accurate understanding of current events and potential developments. His turn towards the social sciences was then largely about finding new ways to task other technical collection systems by cuing off of social findings, e.g. if it was discovered a country's economy was particular sensitive to a particular commodity price, then financial markets could provide indicators of impending political and social instability. JE argued that by extending analysis beyond technical indicators and templates by incorporating social information, customized regional,

national, and subnational models could be developed to provide customized warning for particular cases and improve the quality of warning available to consumers. Importantly, JE noted that while these templates would draw upon social science and social indicators, each would be different and tied to particular cases and circumstances.

JE concluded by observing that one of the difficulties in integrating social science and indicators into the intelligence system was the entrenched institutional interests of expensive, technical collection systems that are the object of budget and corporate priorities. Thus, investments in the social sciences and models faced difficult internal competition when paired against more expensive, traditional collection capabilities, even though the two necessarily enhanced the capabilities and effectiveness of the other.

My next question to JE was about his time as a consumer of intelligence rather than a producer. I was particularly interested in his perspective on what made analysis effective or useful from the perspective of policy-makers. JE answered that he always towards analysts that employed a clear and transparent process for discovering, explaining, and eliminating alternative explanations of events. Whenever analysts did not consider, explore, or provide alternative explanation he noted that he became cautious and worried that analysts were engaged in advocacy for particular views rather than analysis. He also paid great attention to the diversity and use of sources, noting that checks against foreign denial and deception and efforts to cross-validate information were crucial aspects of intelligence and establishing his confidence in the quality of their products. Whenever collection and analysis relied on a single source, he became cautious

and always wanted a way to trace back analytic judgments to the underlying source material.

Finally, I asked JE about the future of intelligence and what he thought would be most important going forward into the future. He noted that our increasing understanding of cognition and the brain was key to improving analysis and understanding the limits, strengths, and failures of human perception and the use of information. He believes that ongoing work in brain scanning and imaging as a way of understanding how beliefs develop and change will make important contributions to improving analysis. Finally, JE noted that understanding cognition will necessarily need to include a better understanding of emotion and its role in analysis and decision making. Indeed, JE believed that many of the difficult intelligence problems are questions of understanding ‘why’ people behave as they do (individually or collectively), and that successfully addressing such concerns requires an understanding of their emotional states, needs, and predisposition as well as more traditional rational considerations.

Interview with Barry Leven, October 25, 2012

Background: Barry Leven (BL) spent more than 28 years working in the intelligence community. His career started with the Navy, where he was a co-op student at a Navy research and development facility in Annapolis, Maryland and worked on technologies for quieting submarines. He then transitioned to Naval Intelligence where he worked on a broad range of acoustic intelligence efforts for detecting submarines and other related foreign military capabilities. He left Naval Intelligence to serve as an Associate National Intelligence Officer for Science and Technology on the National Intelligence Council where he managed the production of National Intelligence Estimates. Upon leaving the National Intelligence Council he joined the Central Intelligence Agency, and at the request of Robert Gates took an assignment in the Strategic Defense Initiative Office (SDIO) providing direct support and managing a small staff drawn from elements of the Intelligence community to General Abrahamson as the DCI's representative. Afterwards, he served as a Division Chief in the CIA's Office of Strategic Weapons and focused on space and strategic systems. He retired from the CIA and the Senior Intelligence Service in 1995.

Discussion: My discussion with BL was far reaching and among the most extensive of those conducted for this project. We spent a great deal of time comparing the cultures of the CIA with naval intelligence. This provided a backdrop for a wide range of topics focused on analytic methods, producer-consumer relations, and more.

My first question to BL asked about his time working in three different intelligence organizations—Naval Intelligence, the CIA, and the NIC. His answers largely provided a context for later questions.

BL's overall sense was that there were very stark differences between the Navy and the CIA. He noted that within Naval Intelligence, career development focused on developing military officers and ensuring their progression within the Navy. In this context, civilians always served as deputies to the military, which simultaneously limited their opportunities to progress while providing greater flexibility for them to pursue their own interests and specialize rather than remain generalists as expected of military officers working on short rotations. Thus, within Naval Intelligence, the leadership rotated through while institutionalized expertise remained embodied in the civilian workforce and enlisted specialists. He noted that many of the senior officers were highly skilled and technically proficient, but their value to the Navy was as generalists rather than specialists, and they were ultimately dependent on their subordinate civilian workforce and enlisted specialists. This dependence created a demanding organizational culture that valued honesty, integrity, and thoroughness out of the mutual dependence between the two classes of employees.

BL also noted that naval intelligence was very close to its consumers, and remained in constant contact with tactical and strategic leadership within the Navy. Intelligence questions were often posed directly and were constantly incoming. Moreover, because the consumers were responsible for the maintenance of the force structure, training, system development and conduct of operations, they cared far more

about the results of intelligence activities as a roadmap for the success of their particular programs and operations. Thus, the emphasis was on outcomes, rather than processes, allowing and encouraging intelligence and operators to work together closely and innovate.

BL noted that the CIA was quite different from the Navy. He argued that ultimately, everything the Navy dealt with was essentially tactical, even when addressing strategic issues such as weapons systems' development and deployment. Questions and responses from the Navy ultimately concerned themselves with ensuring that the Navy's weapons, training, and tactics were appropriately adapted to current and future threats. By contrast, BL noted the CIA consisted of civilians who were much closer in rank and grade to their policy-making consumers, which ironically resulted in greater intellectual and policy distance. As a result, relations between producers and consumers were more difficult at CIA, and policy makers, in his experience, rarely took the time or made the effort to explain their plans and intentions to intelligence analysts, limiting the analysts' ability to meet policy makers' intelligence needs.

BL noted that the great strength of Naval Intelligence was shown in their organizational structures and operations. On the Navy side, personnel, organizational, and budgetary units worked together closely – placing operators, collectors, analysts, and R&D in close proximity and emphasizing their interdependence. He noted that operators, collectors, analysts, and engineers all worked closely and each had direct influence on the activities on their others. Indeed, BL noted that while in naval intelligence a significant

amount of his R&D funding came from system developers reflecting the extent to which intelligence, system development and operational capabilities were integrated.

BL compared the closeness of the Navy with that of the CIA and noted that while the internal organization of the CIA created interdependence between its Directorates, it was not nearly as well-integrated with the organizations it supported as he had observed in the Navy. This was largely because policy makers were further removed from the agency's intelligence activities, and therefore less involved in its production and consumption. He noted that the CIA was a truly executive function while the Navy was more tactical, and that the CIA's interdependencies with policy makers largely focused on ensuring that intelligence provided credible reporting and timely inputs into the policy-process, but was not deeply integrated into high-level deliberations concerning the efficacy of potential courses of action or the appropriateness of policy goals as seen from the intelligence context. By comparison, BL noted that the tight relationship between producers and consumers within the Navy afforded analysts more opportunities to play a stronger role in the deliberations that shaped policy and strategy. For example, he noted that in the Navy it was not uncommon for an Admiral to ask an analyst "what does it mean when you say xyz?" and that an analyst could respond, "Admiral, that isn't the right question." BL noted that such exchanges at the national level between CIA analysts and senior policymakers were far rarer and difficult because policy makers' agendas were more closely held and less likely to be shared with analysts in order to avoid criticism.

Ultimately, BL noted that the stronger institutional integration between intelligence and policy within the Navy when compared with the CIA in the context of

producer and consumer relations created a context where exchanges could be more meaningful, honest, and engaging. Thus, BL felt that the debates between Naval Intelligence and the Navy's senior leadership were more substantive and capable of shifting how each understood the problem and each other's perspectives on it—there was an expectation of a dialog between producers and consumers built upon the trust at senior levels that those beneath them were trying to do the right thing and acted in good faith even when disagreements occurred. By comparison, BL believed that analysts and producer/consumer relations within the CIA was more conservative and restrained when dealing with policy makers, remaining in a position of providing support to consumers and avoiding, whenever possible, challenging their decisions overtly.

BL noted that CIA often sought to develop similar relationships as those found in the Navy and that in theory national intelligence and military intelligence paralleled one another. However, in practice the CIA's environment was very different – essentially playing a smaller role in a larger world given the seniority and diversity of its consumers and distance from them in organizational terms.

Finally, BL noted that the NIC had a different role than the CIA or the Naval Intelligence. During his time, the NIC reported directly to the DCI and was separate from the CIA (it now reports to the DNI). The NIC was composed of analysts and senior members of the academe, government and military drawn from many different disciplines and organizations, and served as the primary interface with policy makers, determining and communicating intelligence priorities to the rest of the community and producing National Intelligence Estimates and other intelligence products. Given its position at the

pinnacle of relations with consumers and diversity of issues and organizations represented by its staff, it drew upon the larger Intelligence Community for intelligence production, and ultimately set the boundaries and focus of production for the rest of the community when dealing with their consumers.

Given the distinct differences between organization, culture, and levels of analysis and engagement with policy makers, I asked BL about the analytic tradecraft of the organizations he worked in and whether it was affected by producer/consumer relations.

BL noted that analytic tradecraft was consistent across organizations, and that despite their differences, analysts approached their problems in similar ways. However, he did note within the Navy, analysts had much narrower accounts and a greater focus on the exploitation of specific INTs. By comparison, analysts in the CIA had broader responsibilities and therefore were far more focused on the integration of multiple INTs or all-source analysis. He noted that time pressures were also relatively consistent across organizations, and that in each case junior analysts were focused on specific tasks and relatively routinized production required by the organization while at senior levels analysts' responsibilities shifted towards answered external requests and producing customized products.

My next question focused on producer/consumer relations and under what conditions consumers viewed analysis as a useful input to shaping policy and when it was dismissed. BL noted that analysts in Naval Intelligence were expected to speak up when confronted with a policymaker who seemed to have misconstrued or misstated Intelligence views, while his time at CIA revealed a much more difficult and tense

relationship with policymakers. He noted that at CIA he was called only rarely into a policy-maker's office to be a sounding board or check on the accuracy of Congressional testimony, while in the Navy it happened regularly and policy makers often changed their testimony as a result of intelligence analysts' inputs. He believed that in the Navy, consumers viewed naval intelligence as partners while at CIA intelligence was seen by consumers as support.

BL noted that relations between producers and consumers usually rested on the strength of the personal relationships between the two. Trust was necessary and highly personal, and while it did not always result in consumers accepting intelligence assessments, it provided an invitation to be a participant in the policy process.

BL also noted that many of the reforms introduced by Robert Gates while at CIA had significant effects on producer/consumer relations. He noted that his placement in the SDIO was indicative of the agency's commitment to improve the relevance of their work to consumers through the development of close personal working relationships with policy makers. He argued that it worked very well, and in the case of the SDIO the development of trust allowed for intelligence analysts to reorient SDIO's assessment of Soviet behavior and attack scenarios in new directions and subsequently change its own plans regarding the development of missile defense systems' capabilities.

Regarding another Gates reform, the insistence that analytic assessments were corporate rather than individual products, BL believed that these changes were less than completely effective in the context of producer/consumer relations even though they

improved the quality and relevance of analysis. The problem was that policy makers placed their trust in individual analysts and preferred to deal with them. They believed that corporate products were unnecessarily watered-down and downplayed important differences between analysts and organizations in order to arrive at an acceptable assessment, masking what were often the most important sources of uncertainty that policy makers were interested in. However, he also noted that policy makers had limited patience and often wanted clear, unambiguous answers and confident predictions, not uncertainty or “on-one-hand and on-the other-hand” reporting which could not always be accommodated given the circumstances of collection, analysis, and the specificity of their concerns.

BL noted that the effectiveness of intelligence was tied to the relevance of the questions that analysts sought to answer. This was achieved by understanding what questions policy makers were asking, and the ways in which they were asking these questions. For example, consumers might rephrase the information found in intelligence assessments, ask for clarification, or challenge judgments, in each case reflecting different motivations.

When asked about the differences between providing intelligence to the executive branch vs. the Congress, BL noted the two were very different kinds of consumers. He noted that the Congress was quite transparent – at least in the House of Representatives – where their interests were almost always tied to concerns about activities in their districts. Thus, threat assessments that justified or challenged the need for particular weapon systems that were manufactured in their district could be of great interest. As a result of

the character of Congressional power and politics, legislative interests in intelligence products were often narrow, critical, accusatory, and pointed. By comparison, the executive branch is responsible for the development and execution of policy, and therefore has different goals and needs from the Congress. Executive intelligence needs and interests were tied to political, organizational, and operational needs of government organizations which were necessarily diverse.

When asked how intelligence avoided becoming politicized, BL argued that there were two key factors. The first was to ensure that everyone in the policy debate received the same intelligence and analytical judgments. Second, was to put policy makers “on the record” by making it known that they received an intelligence assessment. This approach serves to keep the process of giving intelligence to consumers out of the debate by ensuring that everyone worked from the same information (excluding their own personal or organizational sources) and that everyone engaged in the debate knew it.

BL also noted that how information is presented can be part of politicization. He argued that the analytic reasoning process must be transparent to consumers and must be exposed to criticism. The intelligence community can defend its judgments and must advocate for clear and honest descriptions of the intelligence by members of the policy community.

BL also noted that tradecraft matters greatly in producer/consumer relations. He noted that tradecraft included more than just analytic procedures, but also the presentation of information itself and how intelligence collection and judgments are presented. An interesting point, and one that largely echoed points made in prior

interviews with others, was that many observers misunderstand intelligence analysts. Specifically, there is the belief that intelligence analysts are reluctant to share information, but this is incorrect. Analysts have few opportunities to tell people what they know and often relish the opportunity to do so. Instead, they have a tendency to overwhelm consumers with irrelevant information and provide complex assessments when a simpler description or explanation exists. Thus, knowing the consumer's interests and needs is essential, if only to ensure that analysts make appropriate use of their limited time on stage when asked for their inputs.

My next question was focused on the differences between being an analyst and a manager within the IC, and how different responsibilities provided different perspectives on the community. BL noted that as a manager, he developed a much broader perspective on the organization and community. As an analyst, he was worried about his work, but as a manager, he had to think about everyone else's work. He also believed that he was fortunate to have been given a great deal of responsibility for analytic production and management as a young analyst early in his career, and that exposure was important in his professional development. BL noted that as a manager he was constantly focused on the use of time and resources, and engaged in a process of learning how his actions and decisions affected the entire organization for better or worse. He also argued that the key to management was possessing the ability to trust in his subordinates. BL noted as a manager he needed trust his subordinates before they would trust him, thus he needed to "give in order to get," and that his job was ultimately to be vulnerable to the

mistakes of his subordinates while creating an environment that allowed them to do their jobs as best as they could.

When asked about areas that could have worked better, BL noted that collaboration and access both within and outside the intelligence community could have been better and that today's technologies could have enabled much more of it. He stated that while the majority of his career was focused on technical issues regarding foreign weapons capabilities, the ability of new technologies and models to create opportunities to collect and analyze social information about organizations, cultures, social networks, etc. would have all be important additions to their work. BL noted that even in technical areas, the study of social factors was important for understanding the importance of people, places, priorities, and other aspects of foreign military capabilities. Moreover, he noted that huge efforts were required in the past to address these kinds of questions, often taking months to analyze and organize data that could be done in minutes or seconds today.

I asked BL about the particular properties of Agent-Based Modeling (ABM) and the role it could play in analysis. He responded that the ability to integrate political, technical, and economic behavior and processes would have been quite helpful in producing integrated assessments that were traditionally assigned as separate work to different analysts and organizations. Thus, the models themselves could have served as important vehicles within the coordination processes itself, supporting the development of multidisciplinary, all-source, integrated assessments of intelligence targets, e.g. the politics and organization associated with building a ballistic missile capability. Indeed,

BL noted that ABM could have proven quite helpful in the assessment of WMD and missile proliferation, when conventional assessments had little impact getting the community to consider, for example, the role of rogue actors and organizations. In these cases, ABM could have helped identify new collection targets and priorities, allowing for some preliminary evaluations that could subsequently move the system towards recognizing important changes in the system and the need for new approaches and vulnerabilities in order to avoid surprise.

My next question to BL dealt with the problem of prediction and consumer's expectations of analysis. BL acknowledged that consumers always wanted to know what was going to happen and were disappointed when analysts could not provide that information. He noted that analysts never wanted to say that they didn't know what was going to happen, and hated to admit that predictions were not possible. Yet, despite these mutual disappointments, analysis worked best when producers focused on characterizing the range of potential futures and the types of uncertainties that existed regarding what factors or forces might drive outcomes in one direction or another. Often, this occurred by bounding the ranges that particular variables might have, e.g. the number of reentry vehicles the nose-cone of a foreign missile might contain. Such boundaries may be established based on the history of tests, the physical size and capacity of the missile, doctrine, etc. Importantly, uncertainty persists in cases where the best technical intelligence is available because targets can always adapt and change behavior, and often analysts must assess questions about decisions that haven't been made. Again, BL noted in the case of foreign missiles and the number of potential reentry vehicles that the

problem wasn't necessarily estimating the capability to perform a particular technical feat, but rather the desire to do so given the target's other available options and goals.

My final question asked BL what the ideal intelligence community would look like today if he had the opportunity to build one from scratch, unconstrained by existing organizations and history. He argued that he would emphasize people of high integrity and technical skills rather than the structure of organizations themselves. He argued that most analytic shortcomings were not the result of organization, but that analysts did not understand the responsibilities of people they were working with or supporting, and often had little interest in finding out. He also noted that changing the organization of intelligence would matter very little unless the policy community changed its practices and uses of intelligence too. He noted that few policy makers thought deeply about the epistemological limitations and problems of intelligence and intelligence questions. BL argued that greater education on the intelligence community and processes were needed in order to set a better context for producer and consumer relations, and that ultimately, the system would improve when analysts and policy makers could have more frank and open discussions about their responsibilities, interests, and capabilities. He believed that this was the core theme of the reforms that Bob Gates, Doug MacEachin, and other peers and colleagues from his time at CIA were working towards. Moreover, BL argued that the community needed to risk politicization from closeness in order to establish relevance and managers were the best line of defense for defending the integrity of the community.

He concluded by noting the problem with the recruitment system that brings people into the community. BL argued that practices don't work very well and the

descriptions of positions rarely aligned with the real needs of the organization. As a result, the people able to contribute the most over the course their careers were often turned away for not filling a particular niche, and that as long as the positions that were filled by new hires and recruits were narrowly defined, the people best able to handle increasingly broad responsibilities and make diverse commitments were not recruited.

References

References

- Agent-Lab, *Evolution of Central Places (Urban Morphogenesis)*, http://www.maml.hu/maml/model_summaries/EdgeCity/index.html (accessed on September 10, 2012).
- Ahl, Valeria, and T. F. H. Allen, *Hierarchy Theory: A Vision, Vocabulary, and Epistemology* (New York, NY: Columbia University Press, 1996).
- Akerlof, George A., and Rachel E. Kranton, *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being* (Princeton, NJ, Princeton University Press, 2010).
- Alberts, David S., John J. Gartska, Richard E. Hayes, and David A. Signori, *Understanding Information Age Warfare* (Washington, DC: DOD Command and Control Research Project, 2001).
- Andrew, Christopher, *For the President's Eyes Only: Secret Intelligence and the American Presidency from Washington to Bush* (New York, NY: Harper Perennial, 1996).
- Arthur, W. Brian, Steven N. Durlaf and David Lane, eds., *The Economy as an Evolving Complex System II* (Boulder, CO: Westview Press, 1997).
- Axelrod, Robert, *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration* (Princeton, NJ: Princeton University Press, 1997).
- Bailer-Jones, Daniela M., *Scientific Models in Philosophy of Science* (Pittsburgh, PA: University of Pittsburgh Press, 2009).
- Bak, Per, *How Nature Works: The Science of Self-Organized Criticality* (New York, NY: Copernicus, 1996).
- Banks, Steven C., and James J. Gullugly, *Exploratory Modeling: Search Through Spaces of Computational Experiments* (Santa Monica, CA: RAND, 1994).
- Barr, Jason, and Troy Tassier, *Segregation and Strategic Neighborhood Interaction*, Department of Economics, Rutgers University, February 2007,

- <http://andromeda.rutgers.edu/~jmbarr/NYComp/BarrEEA.pdf> (accessed on November 8, 2012).
- Barry, Michael, and Paul M. Torrens, "Modeling Complexity: The Limits to Prediction," *University College of London Working Paper Series*, No. 36 (October 2001).
- Barton, Nicholas H., Derek E.G. Briggs, Jonathan A. Eisen, David B. Goldstein and Nipam H. Paten, *Evolution* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2007).
- Bass, Thomas A., *The Predictors* (New York, NY: Holt, 1999).
- BBC, *Baghdad: Mapping the Violence*,
http://news.bbc.co.uk/2/shared/spl/hi/in_depth/baghdad_navigator/ (accessed on May 8, 2012).
- Beebe, Sarah Miller, *A New Intelligence Technique for Analyzing Complex Adaptive Systems: A Case-Based Assessment*, International Studies Association Annual Conference, March 2011.
- Ben-Israel, Isaac, "Philosophy and Methodology of Intelligence: The Logic of Estimate Process," *Intelligence and National Security*, Vol. 4, No. 4 (1989), pp. 660-718.
- Ben-Israel, Isaac, "Philosophy and Methodology of Military Intelligence: Correspondence with Paul Feyerabend," *Philosophia*, Vol. 28, No. 1-4, (June 2001), pp. 71-102.
- Berg, Nathan, Ulrich Hoffrage and Katarzyna Abramczuk, "Fast Acceptance by Common Experience: FACE-Recognition in Schelling's Model of Neighborhood Segregation," *Judgment and Decision Making*, Vol. 5, No. 5 (August 2010), pp. 391-410.
- Mitchell, Melanie, *An Introduction to Genetic Algorithms* (Cambridge, MA: MIT Press, 1998).
- Bergen, Peter L., *Manhunt: The Ten-Year Search for Bin Laden: From 9/11 to Abbottabad* (New York, NY: Crown, 2012).
- Betts, Richard K., "The New Politics of Intelligence: Will Reforms Work This Time?" *Foreign Affairs*, Vol. 83, No. 3 (May/June, 2004), pp. 2-8.
- Betts, Richard K., *Enemies of Intelligence: Strategic Intelligence and American National Security* (New York, NY: Columbia University Press, 2007).
- Beyerchen, Alan, "Clausewitz, Nonlinearity, and the Unpredictability of War," *International Security*, Vol. 17, No. 3 (Winter 1992-1993), pp. 59-90.

- Biddle, Stephen, "Victory Misunderstood: What the Gulf War Tells Us about the Future of Conflict," *International Security*, Vol. 21, No. 2 (Autumn 1996), pp. 139-179.
- Biddle, Stephen, *Afghanistan and the Future of Warfare: Implications for Army and Defense Policy* (Carlisle, PA: United States Army War College, November 2002).
- Bobbitt, Philip, *The Shield of Achilles: War, Peace, and the Course of History* (New York, NY: Anchor Books, 2002).
- Bracken, Paul, "Net Assessment: A Practical Guide," *Parameters*, Vol. 36, Spring, pp. 90-100 (2006).
- Brewer, Gary D., and Martin Shubik, *The War Game: A Critique of Military Problem Solving* (Cambridge, MA: Harvard University Press, 1979).
- Brown, Chris, and Kristen Ainley, *Understanding International Relations* (New York, NY: Palgrave MacMillan, 2009).
- Brown, Curtis, *Philosophy of Science: Explanation*, January 24, 2011, <http://www.trinity.edu/cbrown/science/explanation.html> (accessed on March 24, 2012).
- Bruce, James B., "Making Analysis More Reliable: Why Epistemology Matters to Intelligence," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 171-190.
- Bruce, James B., "The Missing Link: The Analyst-Collector Relationship," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 191-210.
- Bruce, James B. and Michael Bennett, "Foreign Denial and Deception: Analytical Imperatives," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 122-137.
- Bruce, James B., and Roger Z. George, "Intelligence Analysis—The Emergence of a Discipline" in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence* (Washington, DC: Georgetown University Press, 2008), pp. 1-15.
- Bull, Headley, *The Anarchical Society: A Study of Order in World Politics* (New York, NY: Columbia University Press, 2002).
- Burke, Peter J. and Jan E. Stets, *Identity Theory* (New York, NY: Oxford University Press, 2009).

- Burton, Orville Vernon, ed., *Computing in the Social Sciences and Humanities* (Urbana, IL: University of Illinois Press, 2002).
- Camazine, Scott, Jean-Louis Deneubourg, Nigel R. Franks, James Sneyd, Guy Theraulaz and Eric Bonabeau, *Self-Organization in Biological Systems* (Princeton, NJ: Princeton University Press, 2001).
- Campbell, Joseph Keim, Michael O'Rourke and Matthew H. Slater, eds., *Carving Nature at Its Joints: Natural Kinds in Metaphysics and Science* (Cambridge, MA: MIT Press, 2011).
- Campen, Alan D., ed., *The First Information War* (Fairfax, VA: AFCEA International Press, 1992).
- Carneiro, Robert, "A Theory of the Origin of the State", *Science*, No. 169 (1970), pp. 733-738.
- Cartwright, Nancy, *How the Laws of Physics Lie* (New York, NY: Oxford University Press, 2002).
- Cederman, Lars-Erik, "Rerunning History: Counterfactual Simulation in World Politics," in Philip E. Tetlock and Aaron Belkin, eds., *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives* (Princeton, NJ: Princeton University Press, 1996), pp. 247-267.
- Cederman, Lars-Erik, *Emergent Actors in World Politics: How States and Nations Develop and Dissolve* (Princeton, NJ: Princeton University Press, 1997).
- Central Intelligence Agency, *The Intelligence Cycle*, <https://www.cia.gov/library/publications/additional-publications/the-work-of-a-nation/work-of-the-cia.html> (accessed on June 30, 2012).
- Central Intelligence Agency, *A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 2009).
- Central Intelligence Agency, *Misreading Intentions: Iraq's Reaction to Inspections Created Picture of Deception*, (Washington, DC: Central Intelligence Agency, January 5, 2005), (declassified on June 5, 2012).
- Chermack, Thomas J., *Scenario Planning in Organizations: How to Create, Use, and Assess Scenarios* (San Francisco, CA: Berrett-Koehler Publishers, Inc., 2011).
- Cohen, Eliot A., "Net Assessment: An American Approach," *Jaffe Center for Strategic Studies Memorandum No. 29* (Tel Aviv, IL: Tel Aviv University, 1990).

- Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction, *Report to the President of the United States* (Washington, DC: Government Printing Office, March 31, 2005).
- Congressional Budget Office, *The Costs of Desert Shield* (Washington, DC: Congressional Budget Office, January 1991).
- Creager, Angela N. H., Elizabeth Lunbeck and M. Norton Wise, "Introduction," in Angela N. H. Creager, Elizabeth Lunbeck and M. Norton Wise, eds., *Science Without Laws: Model Systems, Cases, Exemplary Narratives* (Durham, NC: Duke University Press, 2007), pp. 1-32.
- Creager, Angela N. H., Elizabeth Lunbeck and M. Norton Wise, eds., *Science Without Laws: Model Systems, Cases, Exemplary Narratives* (Durham, NC: Duke University Press, 2007).
- Cronin, Audry Kurth, *How Terrorism Ends: Understanding the Decline and Demise of Terrorist Campaigns* (Princeton, NJ: Princeton University Press, 2009).
- Crosby, Alfred W., *The Measure of Reality: Quantification and Western Society, 1250-1600* (New York, NY: Cambridge University Press, 1997).
- Crowther-Heyck, Hunter, *Herbert A. Simon: The Bounds of Reason in Modern America* (Baltimore, MD: Johns Hopkins University Press, 2005).
- Curd, Martin, and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton and Company, 1998).
- Cusack, Thomas R., and Richard J. Stoll, *Exploring Realpolitik: Probing International Relations Theory with Computer Simulation* (Boulder, CO: Lynne Rienner Publishers, 1990).
- d'Inverno, Mark, and Michael Luck, *Understanding Agent Systems* (New York, NY: Springer, 1998).
- Danzig, Richard, *Driving in the Dark: Ten Propositions About Prediction and National Security* (Washington, DC: Center for New American Security, October 2011).
- Darley, Vincent, and Alexander V. Outkin, *A NASDAQ Market Simulation: Insights on a Major Market From the Science of Complex Adaptive Systems* (Hackensack, NJ: World Scientific, 2007).
- Davis, Jack, "A Policymaker's Perspective On Intelligence Analysis," *Studies in Intelligence*, Vol. 38, No. 5 (1995), pp. 7-15.

- Davis, Jack, *A Compendium of Tradecraft Notes* (Washington, DC: Central Intelligence Agency, February 1997).
- Davis, Jack, "Why Bad Things Happen to Good Analysts," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 157-170.
- De Jong, Ken, *Evolutionary Computation: A Unified Approach* (Cambridge, MA: MIT Press, 2002).
- Der Derian, James, "Anti-Diplomacy, Intelligence Theory and Surveillance Practice," *Intelligence and National Security*, Vol. 8, No. 3 (July 1993), pp. 29-51.
- Dokumaci, Emin, and William H. Sandholm, Schelling Redux: An Evolutionary Dynamic Model of Residential Segregation, Department of Economics, University of Wisconsin, June 19, 2007, <http://www.ssc.wisc.edu/~whs/research/schelling.pdf> (accessed on November 8, 2012).
- Doran, Jim, and Nigel Gilbert, "Simulating Societies: An Introduction," in Nigel Gilbert and Jim Doran, eds., *Simulating Societies: The Computer Simulation of Social Phenomena* (London, UK: University College of London Press, 1994), pp. 1-18.
- Dorner, Dietrich, *The Logic of Failure: Recognizing and Avoiding Error in Complex Situations* (Cambridge, MA: Perseus Books, 1996).
- Dougherty, James E., and Robert L. Pfaltzgraff, Jr., *Contenting Theories of International Relations: A Comprehensive Survey* (New York, NY: Pearson, 2009).
- Doyle, Michael W., and G. John Ikenberry, eds., *New Thinking in International Relations Theory* (Boulder, CO: Westview Press, 1997).
- Edwards, Sean J., *Swarming and the Future of Warfare* (Santa Monica, CA: RAND, 2004).
- Eickelman, Dale F., and James Piscatori, *Muslim Politics* (Princeton, NJ: Princeton University Press, 2004).
- Elman, Colin, and Miriam Fendius Elman, eds., *Bridges and Boundaries: Historians, Political Scientists, and the Study of International Relations* (Cambridge, MA: MIT Press, 2001).
- Epstein, Joshua M., *Generative Social Science: Studies in Agent-Based Computational Modeling* (Princeton, NJ: Princeton University Press, 2006).

- Epstein, Joshua M., and Robert Axtell, *Growing Artificial Societies: Social Science from the Bottom Up* (Cambridge, MA: MIT Press 1996).
- Finager, Thomas, "Office of the Director of National Intelligence: Promising Start Despite Ambiguity, Ambivalence, and Animosity," in Roger Z. George and Harvey Rishikof, eds., *The National Security Enterprise: Navigating the Labyrinth* (Washington, DC: Georgetown University Press, 2011), pp. 139-155.
- Fingar, Thomas, *Reducing Intelligence: Intelligence Analysis and National Security* (Stanford, CA: Stanford University Press, 2011).
- Flannery, Kent V., "The Cultural Evolution of Civilizations," *Annual Review of Ecology and Systematics*, Vol. 3 (1972), pp. 399-426.
- Flyvbjerg, Bent, *Making Social Science Matter: Why Social Inquiry Fails and How it Can Succeed Again* (New York, NY: Cambridge University Press, 2001).
- Frank, Aaron B., "Pre-Conflict Management Tools: Winning the Peace," *Defense and Technology Paper, No. 11* (Washington, DC: Center for Technology and National Security Policy, National Defense University, February 2005).
- Frank, Aaron B. *Interview with Paul Pillar from Georgetown University*, <http://www.aaronbfrank.com/2012/06/interview-with-paul-pillar-from-georgetown-university/> (accessed on June 19, 2012).
- Fujita, Mashahisa, Paul Krugman, and Anthony J. Venables, *The Spatial Economy* (Cambridge, MA: MIT Press, 2001).
- Fukuyama, Frances, ed., *Blindside: How to Anticipate Forcing Events and Wild Cards in Global Politics* (Washington, DC: Brookings, 2007).
- Fuller, Graham E., and Rend Rahim Francke, *The Arab Shi'a* (New York, NY: St. Martin's Press, 1999).
- Gaddis, John Lewis, *The United States and the End of the Cold War: Implications, Reconsiderations, Provocations* (New York, NY: Oxford University Press, 1994).
- Gaddis, John Lewis, *We Now Know: Rethinking Cold War History* (New York, NY: Oxford University Press, 1998).
- Gaddis, John Lewis, "In Defense of Particular Generalization: Rewriting Cold War History, Rethinking International Relations Theory," in Colin Elman and Miriam Fendius Elman, eds., *Bridges and Boundaries: Historians, Political Scientists, and the Study of International Relations* (Cambridge, MA: MIT Press, 2001), pp. 301-326.

- Gaddis, John Lewis, *The Landscape of History: How Historians Map the Past* (New York, NY: Oxford University Press, 2002).
- Gat, Azar, *War and Human Civilization* (New York, NY: Oxford University Press, 2008).
- Gates, Robert M., "Guarding Against Politicization," *Studies in Intelligence*, Vol. 36, No 1 (1992), pp. 5-13.
- George, Alexander L., *Bridging the Gap: Theory and Practice in Foreign Policy* (Washington, DC: United States Institute of Peace, 1993).
- George, Roger Z., "Fixing the Problem of Analytical Mind-Sets," *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), pp. 385-404.
- George, Roger Z., "The Art of Strategy and Intelligence" in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 107-137.
- George, Roger Z., "Reflections on CIA Analysis: Is it Finished?" *Intelligence and National Security*, Vol. 26, No. 1 (February 2011), pp. 72-81.
- George, Roger Z., "Central Intelligence Agency: The President's Own," in Roger Z. George and Harvey Rishikof, eds., *The National Security Enterprise: Navigating the Labyrinth* (Washington, DC: Georgetown University Press, 2011), pp. 157-173.
- George, Roger Z., and Harvey Rishikof, eds., *The National Security Enterprise: Navigating the Labyrinth* (Washington, DC: Georgetown University Press, 2011).
- George, Roger Z., and James B. Bruce, "Glossary of Analysis Terms," in Roger Z. George and James B. Bruce, *Analyzing Intelligence: Origins, Obstacles, Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 309-319.
- George, Roger Z., and James B. Bruce, eds., *Analyzing Intelligence* (Washington, DC: Georgetown University Press, 2008).
- Ghani, Ashraf, and Clare Lockhart, *Fixing Failed States: A Framework for Rebuilding a Fractured World* (New York, NY: Oxford University Press, 2009).
- Gilbert, Nigel and Rosaria Conte, eds., *Artificial Societies: the Computer Simulation of Social Life* (New York, NY: Routledge, 1995).
- Gilbert, Nigel, and Jim Doran, eds., *Simulating Societies: The Computer Simulation of Social Phenomena* (London, UK: University College of London Press, 1994).

- Gill, Jeff, *Essential Mathematics for Political and Social Research* (New York, NY: Cambridge University Press, 2006).
- Gill, Peter, and Mark Phythian, *Intelligence in an Insecure World* (Malden, MA: Polity, 2008).
- Gleick, James, *Chaos* (New York, NY: Penguin Books, 1987).
- Gode, Dhananjay K. and Shyam Sunder, "Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality," *The Journal of Political Economy*, Vol. 101, No. 1 (February 1993), pp. 119-137.
- Godfrey-Smith, Peter, *Theory and Reality: An Introduction to the Philosophy of Science* (Chicago, IL: University of Chicago Press, 2003).
- Godfrey-Smith, Peter, "The Strategy of Model-Based Science," *Biology and Philosophy*, Vol. 21, (2006), pp. 725-740.
- Godfrey-Smith, Peter, "Induction, Samples, and Kinds," in Joseph Keim Campbell, Michael O'Rourke and Matthew H. Slater, eds., *Carving Nature at Its Joints: Natural Kinds in Metaphysics and Science* (Cambridge, MA: MIT Press, 2011), pp. 33-51.
- Godson, Roy, Ernest R. May and Gary Schmitt, eds., *U.S. Intelligence at the Crossroads: Agendas for Reform* (Washington, DC: Brassey's, 1995).
- Grant, Peter R. and B. Rosemary Grant, *How and Why Species Multiply: The Radiation of Darwin's Finches* (Princeton, NJ: Princeton University Press, 2008).
- Grimm, Volker and Patrick Railsback, *Individual-Based Modeling and Ecology* (Princeton, NJ: Princeton University Press, 2005).
- Groves, David G. and Robert J. Lempert, "A New Analytic Method for Finding Policy-Relevant Scenarios" *Global Environmental Change*, Vol.17, No. 1 (2007), pp. 73-85.
- Guilaine, Jean and Jean Zammit, *The Origins of War: Violence in Prehistory* (Malden, MA, Blackwell Publishing, 2005).
- Haftendorn, Helga, "The Security Puzzle: Theory-Building and Discipline-Building in International Security," *International Studies Quarterly*, Vol. 35, No. 1 (March 1991), pp. 3-17.
- Hall, Wayne Michael, and Gary Citrenbaum, *Intelligence Analysis: How to Think in Complex Environments* (Santa Barbara, CA: Praeger Security International, 2010).

- Handel, Michael H., "Strategic Surprise: The Politics of Intelligence and the Management of Uncertainty," in Alfred C. Mauer, Marion D. Tunstall and James M. Keagle, eds., *Intelligence: Policy and Process* (Boulder, CO: Westview Press, 1985), pp. 239-269.
- Hawkins, Ed, *The Importance of Reliable Uncertainty Estimates*, January 3, 2012, <http://www.climate-lab-book.ac.uk/2012/reliable-uncertainty/> (accessed on June 20, 2012).
- Hedley, John H., "The Evolution of Intelligence Analysis," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 19-34.
- Hedley, John Hollister, "Analysis for Strategic Intelligence," in Loch K. Johnson, ed., *Handbook of Intelligence Studies* (New York, NY: Routledge, 2009), pp. 211-226.
- Helfrick, Albert, *Electronics in the Evolution of Flight* (College Station, TX: Texas A&M Press, 2004).
- Herman, Mark, "Entropy-Based Warfare: Modeling the Revolution in Military Affairs," *Joint Forces Quarterly*, No. 40 (Winter 1998-99), pp. 85-90.
- Herman, Mark, Mark Frost, and Robert Kurz, *Wargaming for Leaders: Strategic Decision Making from the Battlefield to the Boardroom* (New York, NY: McGraw Hill, 2009).
- Heuer, Richards J., Jr., "Adapting Academic Methods and Models to Government Needs," in Richards J. Heuer, Jr., ed., *Quantitative Approaches to Political Intelligence: The CIA Experience* (Boulder, CO: Westview Press, 1978), pp. 1-10.
- Heuer, Richards J., Jr., ed., *Quantitative Approaches to Political Intelligence: The CIA Experience* (Boulder, CO: Westview Press, 1978).
- Heuer, Richards J., Jr., *Psychology of Intelligence Analysis* (Washington, DC: Central Intelligence Agency, 1999).
- Heuer, Richards J., Jr., "Computer-Aided Analysis of Competing Hypotheses," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 251-265.
- Heuer, Richards J., Jr., and Randolph Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: Congressional Quarterly, 2011).

- Hilsman, Roger, Jr., "Intelligence and Policy-Making in Foreign Affairs," *World Politics*, Vol. 5, No. 1 (October 1952), pp. 1-45.
- Hodges, James S., "Six (Or So) Things You Can Do with a Bad Model," *Operations Research*, Vol. 39, No. 3 (May-June, 1991), pp. 355-365.
- Hoffman, Bruce, "All You Need is Love: How the Terrorists Stopped Terrorism," *The Atlantic Monthly*, Vol. 288, No. 5 (December 2001), pp. 34-37.
- Horowitz, Irving Louis, ed., *The Rise and Fall of Project Camelot* (Cambridge, MA: MIT Press, 1976).
- Howard, Michael, *War and European History* (New York, NY: Oxford University Press, 2009).
- Hulnick, Arthur S., "The Intelligence Cycle," in Loch K. Johnson and James Wirtz, eds., *Intelligence: The Secret World of Spies* (New York, NY: Oxford University Press, 2011), pp. 65-77.
- Huntington, Samuel P., *The Clash of Civilizations and the Remaking of World Order* (New York, NY: Touchstone, 1996).
- Ikenberry, G. John, *After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order After Major Wars* (Princeton, NJ: Princeton University Press, 2001).
- Ilachinski, Andrew, *Artificial War: Multiagent-Based Simulation of Combat* (River Edge, NJ: World Scientific, 2004).
- Institute for Defense Analysis, *Net Assessment: The Concept, Its Development and Its Future* (Alexandria, VA: Institute for Defense Analysis, May 22, 1990).
- Jenkins, Richard, *Social Identity* (New York, NY: Routledge, 2008).
- Jervis, Robert, *The Logic of Images in International Relations* (New York, NY: Columbia University Press, 1970).
- Jervis, Robert, "Improving the Intelligence Process: Informal Norms and Incentives," in Alfred C. Maurer, Marion D. Tunstall and James M. Keagle, eds., *Intelligence: Policy and Process* (Boulder, CO: Westview Press, 1985), pp. 113-124.
- Jervis, Robert, *Why Intelligence Fails: Lessons from the Iranian Revolution and the Iraq War* (Ithaca, NY: Cornell University Press, 2010).
- Johnson, Loch K., "Introduction," in Loch K. Johnson, ed., *Handbook of Intelligence Studies* (New York, NY: Taylor and Francis, 2009), pp. 1-14.

- Johnson, Loch K., ed., *Handbook of Intelligence Studies* (New York, NY: Taylor and Francis, 2009).
- Johnson, Loch K., and James Wirtz, eds., *Intelligence: The Secret World of Spies* (New York, NY: Oxford University Press, 2011).
- Johnson, Loch K., ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010).
- Joint Chiefs of Staff, *Joint Publication 5-0: Joint Planning Operations* (Washington, DC: Department of Defense, August 11, 2011).
- Jones, Morgan D., *The Thinkers Toolkit: 14 Powerful Techniques for Problem Solving* (New York, NY: Crown, 1998).
- Kapstein, Ethan B., and Michael Mastanduno, eds., *Unipolar Politics: Realism and State Strategies After the Cold War* (New York, NY: Columbia University Press, 1999).
- Keaney, Thomas A. and Eliot A. Cohen, *Revolution in Warfare? Air Power in the Persian Gulf* (Annapolis, MD: Naval Institute Press, 1995).
- Kendall, Willmoore, "The Function of Intelligence," *World Politics*, Vol. 1, No. 4 (July 1949), pp. 542-552.
- Kent Center, "Rethinking "Alternative Analysis" to Address Transnational Threats," *Occasional Papers*, Vol. 3, No. 2 (October 2004).
- Kent, Sherman, "Estimates and Influence," *Studies in Intelligence*, Vol. 12, No. 3 (1968), pp. 11-21.
- Kent, Sherman, *Strategic Intelligence for American World Policy* (Princeton, NJ: Princeton University Press, 1949).
- Keohane, Robert O., *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton, NJ: Princeton University Press, 2005).
- Kerbel, Joshua and Anthony Olcott, "Synthesizing with Clients, Not Analyzing for Customers," *Studies in Intelligence*, Vol. 54, No. 4 (December 2010), pp. 11-27.
- Khaneman, Daniel, *Thinking Fast and Slow* (New York, NY: Farrar, Straus and Giroux, 2011).
- Kilcullen, David, *The Accidental Guerrilla: Fighting Small Wars in the Midst of a Big One* (New York, NY: Oxford University Press, 2009).

- Klahr, David, *Exploring Science: The Cognition and Development of Discovery Processes* (Cambridge, MA: MIT Press, 2002).
- Klemke, E. D., "What is Philosophy of Science?" in E. D. Klemke, Robert Hollinger and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), pp. 19-37.
- Klemke, E. D., Robert Hollinger and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998).
- Klien, Gary, *Sources of Power: How People Make Decisions* (Cambridge, MA: MIT Press, 1998).
- Knusten, Torbjorn L., *A History of International Relations Theory* (New York, NY: Manchester University Press, 1997).
- Kohler, Timothy A., and George J. Gumerman, eds., *Dynamics in Human and Primate Societies: Agent-Based Modeling of Social and Spatial Processes* (New York, NY: Oxford University Press, 2000).
- Kovacs, Amos, "The Nonuse of Intelligence," *International Journal of Intelligence and Counterintelligence*, Vol. 10, No. 4 (December 1997), pp. 383-417.
- Krepinevich, Andrew F., Jr., *The Military-Technical Revolution: A Preliminary Assessment* (Washington, DC: Center for Strategic and Budgetary Assessments, 2002).
- Krepinevich, Andrew F., *7 Deadly Scenarios: A Military Futurist Explores War in the 21st Century* (New York, NY: Bantam, 2009).
- Kries, John F., ed., *Piercing the Fog: Intelligence and Army Air Forces Operations in World War II* (Washington, DC: Air Force History and Museums Program, 1995).
- Krugman, Paul, *The Self-Organizing Economy* (Malden, MA: Blackwell, 1996).
- Krugman, Paul, "How the Economy Organizes Itself in Space: A Survey of New Economic Geography," in W. Brian Arthur, Steven N. Durlaf, and David A. Lane, eds., *The Economy as a Complex Adaptive System II* (Boulder, CO: Westview Press, 1997), pp. 239-262.
- Kuhn, Thomas S., "Logic of Discovery or Psychology of Research?" in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton and Company, 1998), pp. 11-19.

- Lacquer, Walter, *A World of Secrets: The Uses and Limits of Intelligence* (New York, NY: Basic Books, 1985).
- Laird, John E., *The SOAR Cognitive Architecture* (Cambridge, MA: MIT Press, 2012).
- Lakatos, Imre, "Science and Pseudoscience," in Martin Curd and J. A. Cover, eds., *Philosophy of Science: The Central Issues* (New York, NY: W. W. Norton and Company, 1998), pp. 20-26.
- Langley, Pat, Herbert A. Simon, Gary L. Bradshaw and Jan M. Zytkow, *Scientific Discovery: Computational Explorations of the Creative Process* (Cambridge, MA: MIT Press, 1987).
- Lave, Charles A., and James G. March, *An Introduction to Models in the Social Sciences* (Lanham, MD: University Press of America, 1993).
- Laver, Michael, and Ernest Sergenti, *Party Competition: An Agent-Based Model* (Princeton, NJ: Princeton University Press, 2012).
- Lawler, Steph, *Identity: Sociological Perspectives* (New York, NY: Policy, 2009).
- Lebow, Richard Ned, and Thomas Risse-Kappen, eds., *International Relations Theory at the End of the Cold War* (New York, NY: Columbia University Press, 1994).
- Lempert, Robert, "Can Scenarios Help Policymakers Be Both Bold and Careful?" in Frances Fukuyama, ed., *Blindside: How to Anticipate Forcing Events and Wild Cards in Global Politics* (Washington, DC: Brookings, 2007), pp. 109-119.
- Lempert, Robert J., David G. Groves, Steven W. Popper, and Steve C. Bankes, "A General, Analytic Method for Generating Robust Strategies and Narrative Scenarios," *Management Science*, Vol. 52, No. 4 (April 2006), pp. 514-528.
- Lempert, Robert J., Steven W. Popper, and Steven C. Bankes, *Shaping the Next One Hundred Years: New Methods for Long-Term Quantitative Policy Analysis* (Santa Monica, CA: RAND, 2003).
- Levin, Simon, *Fragile Dominion: Complexity and the Commons* (Cambridge, MA: Perseus Publishing, 1999).
- Levin, Simon A., ed., *The Princeton Guide to Ecology* (Princeton, NJ: Princeton University Press, 2009).
- Lewis, Michael, "Obama's Way," *Vanity Fair*, October 2012, <http://www.vanityfair.com/politics/2012/10/michael-lewis-profile-barack-obama> (accessed on September 17, 2012).

- Lindgren, Mats and Hans Bandhold, *Scenario Planning: The Link Between Future and Strategy* (New York, NY: Palgrave Macmillan, 2009).
- Looney, Robert E., "DARPA's Policy Analysis Market for Intelligence: Outside the Box or Off the Wall?" *International Journal of Intelligence and Counterintelligence*, Vol. 17, No. 3 (2004), pp. 405-419.
- Lowenthal, Mark M., "Tribal Tongues: Intelligence Consumers, Intelligence Producers," *The Washington Quarterly*, Vol. 15, No. 1 (1992), pp. 157-168.
- Lowenthal, Mark M., "Intelligence Analysis: Management and Transformation Issues," in Jennifer E. Simms and Burton Gerber, eds., *Transforming U.S. Intelligence* (Washington, DC: Georgetown University Press, 2005), pp. 220-2238.
- Lowenthal, Mark M., "Intelligence in Transition: Analysis after September 11 and Iraq," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 226-237.
- Lowenthal, Mark M., "Towards a Reasonable Standard for Analysis: How Right, How Often on Which Issues?" *Intelligence and National Security*, Vol. 23, No. 3 (2008), pp. 303-315.
- Lowenthal, Mark M., *Intelligence: From Secrets to Policy* (Washington, DC: Congressional Quarterly Press, 2009).
- Lowenthal, Mark M., "The Policymaker-Intelligence Relationship," in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), pp. 437-451.
- Lustick, Ian S., Dan Miodownik, and Roy J. Eidelson, "Secessionism in Multicultural States: Does Sharing Power Prevent or Encourage It?" *American Political Science Review*, Vol. 98, No. 2 (May 2004), pp. 209-229.
- Luttwak, Edward N., *Strategy: The Logic of War and Peace* (Cambridge, MA: The Belknap Press of Harvard University Press, 2001).
- MacEachin, Douglas J., "The Tradecraft of Analysis," in Roy Godson, Ernest R. May and Gary Schmitt, eds., *U.S. Intelligence at the Crossroads Agendas for Reform* (Washington, DC: Brassey's, 1995), pp. 63-75.
- MacEachin, Douglas, "Analysis and Estimates: Professional Practices in Intelligence Production," in Jennifer E. Sims and Burton Gerber, eds., *Transforming U.S. Intelligence* (Washington, DC: Georgetown University Press, 2005), pp. 115-133.

- Mahnken, Thomas G., *Uncovering Ways of War: U.S. Intelligence and Foreign Military Innovation, 1918-1941* (Ithaca, NY: Cornell University Press, 2002).
- Mann, James, *Rise of the Vulcans: The History of Bush's War Cabinet* (New York, NY: Viking, 2004).
- March, James G., "Making Artists Out of Pedants," in Ralph M. Stogdill, ed., *The Process of Modeling-Building in the Social Sciences* (New York, NY: W.W. Norton and Company, 1970), pp. 54-75.
- March, James G., *Decisions and Organizations* (New York, NY: Basic Books, 1988).
- Marcus, Alfred, *Strategic Foresight: A New Look at Scenarios* (New York, NY: Palgrave Macmillan, 2009).
- Marrin, Stephen, "Is Intelligence Analysis an Art or a Science?" *International Journal of Intelligence and Counterintelligence*, Vo. 25, No. 3 (2012), pp. 529-545.
- Marshall, Andrew W., J. J. Martin, and Henry S. Rowen, eds., *On Not Confusing Ourselves: Essays on National Security Strategy in Honor of Albert and Roberta Wohlstetter* (Boulder, CO: Westview Press, 1991).
- Maurer, Alfred C., Marion D. Tunstall and James M. Keagle, eds., *Intelligence: Policy and Process* (Boulder, CO: Westview Press, 1985).
- May, Ernst R., ed., *Knowing One's Enemies: Intelligence Assessment Before the Two World Wars* (Princeton, NJ: Princeton University Press, 1986).
- McAdam, Doug, Sidney Tarrow and Charles Tilly, *Dynamics of Contention* (New York, NY: Cambridge University Press, 2001).
- McDowell, Don, *Strategic Intelligence: A Handbook for Practitioners, Managers, and Users* (Lanham, MD: Scarecrow Press, 2009).
- McLaughlin, John, "Serving the National Policymaker," in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 71-81.
- McNeill, William H., *The Pursuit of Power* (Chicago, IL: University of Chicago Press, 1982).
- McNeill, William H., *The Rise of the West: A History of Human Community* (Chicago, IL: University of Chicago Press, 1991).
- Mearsheimer, John J., *The Tragedy of Great Power Politics* (New York, NY: W. W. Norton and Company, 2001).

- Miller, John H., and Scott E. Page, *Complex Adaptive Systems: An Introduction to Computational Models of Social Life* (Princeton, NJ: Princeton University Press, 2007).
- Mitchell, Gordon R., "Team B Intelligence Coups," *Quarterly Journal of Speech*, Vol. 92, No. 2 (2006), pp. 144-173.
- Mitchell, Melanie, *An Introduction to Genetic Algorithms* (Cambridge, MA: MIT Press, 1996).
- Montgomery, Evan Braden, *Defense Planning for the Long Haul: Scenarios, Operational Concepts, and the Future Security Environment* (Washington, DC: Center for Strategic and Budgetary Assessments, 2009).
- Morgan, Mary S., and Margaret Morrison, eds., *Models as Mediators: Perspectives on Natural and Social Science* (New York, NY: Cambridge University Press, 1999).
- Morganthau, Hans J., *Politics Among Nations: The Struggle for Power and Peace* (Boston, MA: McGraw-Hill, 1993).
- Morrison, Margaret, "Models as Autonomous Agents," in Mary S. Morgan and Margaret Morrison, eds., *Models as Mediators: Perspectives on Natural and Social Science* (New York, NY: Cambridge University Press, 1999), pp. 38-65.
- Moynihan, Daniel Patrick, *Secrecy* (New Haven, CT: Yale University Press, 1998).
- Mueller, John, "The Essential Irrelevance of Nuclear Weapons," *International Security*, Vol. 13, No. 2 (Autumn 1988), p. 55-79.
- Mueller, John, *Quiet Cataclysm: Reflections of the Recent Transformation of World Politics* (New York, NY: Harper Collins, 1995).
- Murray, Williamson, and Allan R. Millett, eds., *Calculations: Net Assessment and the Coming of World War II* (New York, NY: The Free Press, 1992).
- Nakash, Yitzhak, *Reaching for Power* (Princeton, NJ: Princeton University Press, 2006).
- Nasr, Vali, *The Shia Revival: How Politics within Islam Will Shape the Future* (New York, NY: W. W. Norton and Company, 2006).
- National Intelligence Council, *Global Trends 2010*
http://www.dni.gov/nic/special_globaltrends2010.html (accessed on June 17, 2012).

- National Intelligence Council, *Global Trends 2015: A Dialogue About the Future With Nongovernment Experts* (Washington, DC: National Intelligence Council, December 2000).
- National Intelligence Council, *Global Trends 2025: A Transformed World* (Washington, DC: Directorate of National Intelligence, November 2008).
- National Intelligence Council, *Mapping the Global Future: Report on the National Intelligence Council's 2020 Project Based on Consultations With Nongovernmental Experts Around the World* (Washington, DC: National Intelligence Council, December 2004).
- Naval Historical Records Center, "USS *Albacore* (AGSS-569), 1953-1984," <http://www.history.navy.mil/photos/sh-usn/usnsh-a/ss569.htm> (accessed on August 5, 2011).
- Neustadt, Richard E., and Ernst R. May, *Thinking in Time: The Uses of History for Decision Makers* (New York, NY: The Free Press, 1986).
- Nielsen, Michael, *Reinventing Discovery: The New Era of Networked Science* (Princeton, NJ: Princeton University Press, 2011).
- Ninic, Miroslav, and Joseph Leggold, eds., *Being Useful: Policy relevance and International Relations Theory* (Ann Arbor, MI: University of Michigan Press, 2000).
- Nolte, William M., "Intelligence Analysis in an Uncertain Environment," in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), pp. 404-421.
- Nye, Joseph S., Jr., and Roger K. Smith, eds., *After the Storm: Lessons From the Gulf War* (Lanham, MD: Madison Books, 1992).
- Nye, Joseph S., Jr., and William A. Owens, "America's Information Edge," *Foreign Affairs*, Vol. 75, No. 2 (March/April 1996), pp. 20-36.
- O'Hanlon, Michael, "Estimating Casualties in a War to Overthrow Saddam," *Orbis* (Winter 2003), pp. 21-40.
- O'Leary, Michael K., William D. Coplin, Howard B. Shapiro and Dale Dean, "The Quest for Relevance: Quantitative International Relations Research and Government Foreign Affairs Analysis," *International Studies Quarterly*, Vol. 18, No. 2 (June 1974), pp. 211-237.

- O'Toole, G. J. A., *Honorable Treachery: A History of U.S. Intelligence, Espionage, and Covert Action from the American Revolution to the CIA* (New York, NY: Atlantic Monthly Press, 1993).
- Ogilvy, James A., *Creating Better Futures: Scenario Planning as a Tool for a Better Tomorrow* (New York, NY: Oxford University Press, 2002).
- Okasha, Samir, *Philosophy of Science: A Very Short Introduction* (New York, NY: Oxford University Press, 2002).
- Online Philosophy Club, *Philosophy of Science*, <http://onlinephilosophyclub.com/philosophy-of-science.php> (accessed on January 26, 2012).
- Overbye, Dennis, "A Scientist Takes On Gravity," *The New York Times*, July 12, 2010, available at: <http://www.nytimes.com/2010/07/13/science/13gravity.html?pagewanted=all> (accessed on March 24, 2012).
- Owens, Bill, with Ed Offley, *Lifting the Fog of War* (New York, NY: Farrar, Straus and Giroux, 2000).
- Owens, William A., "JROC: Harnessing the Revolution in Military Affairs," *Joint Forces Quarterly*, No. 5 (Summer 1994), pp. 55-57.
- Pacey, Arnold, *Technology in World Civilization: A Thousand-Year History* (Cambridge, MA: MIT Press, 1991).
- Palmer, Mark, *Gain Capital: A Glimpse in the Future of Analytics?*, May 4, 2012, http://streambase.typepad.com/streambase_stream_process/2012/05/inside-gain-capitals-use-of-streambase-liveview.html (accessed on June 23, 2012).
- Parker, Geoffrey, *The Military Revolution: Military Innovation and the Rise of the West, 1500-1800* (New York, NY: Cambridge University Press, 1996).
- Parunak, H. Van Dyke, Robert Savit, and Rick L. Riolo, "Agent Based Modeling Vs. Equation-Based Modeling: A Case Study and User's Guide," *Proceedings of Workshop on Modeling Agent Based Systems*, Paris, 1998.
- Perrow, Charles, *Complex Organizations: A Critical Essay* (New York, NY: McGraw Hill, 1986).
- Perry, William J., "Desert Storm and Deterrence in the Future," in Joseph S. Nye, Jr. and Roger K. Smith, eds., *After the Storm: Lessons From the Gulf War* (Lanham, MD: Madison Books, 1992), pp. 241-263.

- Petersen, Martin, "What We Should Demand from Intelligence," *National Security Studies Quarterly* Vol. 5, No. 2 (Spring 1999), pp. 107-114.
- Petersen, Troels, *Measuring the W Mass with ATLAS*, February 11, 2008, http://atlas-service-enews.web.cern.ch/atlas-service-enews/2007-8/features_07-8/features_wmass.php (accessed on June 20, 2012).
- Pherson Associates, "Analytic Tools and Techniques" at <http://www.pherson.org/Tools.html> (accessed on June 16, 2012).
- Pickett, George E., James G. Roche, and Barry D. Watts, "Net Assessment: A Historical Review," in Andrew W. Marshall, J. J. Martin, and Henry S. Rowen, eds., *On Not Confusing Ourselves: Essays on National Security Strategy in Honor of Albert and Roberta Wohlstetter* (Boulder, CO: Westview Press, 1991), pp. 158-185.
- Pillar, Paul R., "Intelligence, Policy, and the War in Iraq," *Foreign Affairs*, Vol. 85, No. 2 (March/April 2006), pp. 15-27.
- Pillar, Paul R., "The Right Stuff," *The National Interest*, No. 91 (September/October 2007), pp. 53-59.
- Pillar, Paul R., "The Perils of Politicization," in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), pp. 472-484.
- Pillar, Paul R., *Intelligence and U.S. Foreign Policy: Iraq, 9/11, and Misguided Reform* (New York, NY: Columbia University Press, 2011).
- Pomeranz, Kenneth, *The Great Divergence: China, Europe, and the Making of the Modern World Economy* (Princeton, NJ: Princeton University Press, 2000).
- Popper, Karl, *The Poverty of Historicism* (New York, NY: Routledge, 2004).
- Popper, Karl, *Conjectures and Refutations* (New York, NY: Routledge, 2010).
- Posner, Richard A., *Uncertain Shield: The U.S. Intelligence in the Throes of Reform* (Lanham, MD: Rowman and Littlefield, 2006).
- Post, Jerrold M., *The Mind of the Terrorist: The Psychology of Terrorism from the IRA to Al-Qaeda* (New York, NY: Palgrave, 2007).
- Prunckun, Hank, *Handbook of Scientific Methods of Inquiry for Intelligence Analysis* (Lanham, MD: Scarecrow Press, 2010).
- Putnam, Hilary, *The Collapse of the Fact/Value Dichotomy and Other Essays* (Cambridge, MA: Harvard University Press, 2002).

- Reynolds, Craig, *Boids: Background and Update*, <http://www.red3d.com/cwr/boids/> (accessed on August 7, 2011).
- Rice, Condoleezza, “Why We Know Iraq is Lying,” *The New York Times*, January 23, 2003, <http://www.nytimes.com/2003/01/23/opinion/why-we-know-iraq-is-lying.html?pagewanted=all&src=pm> (accessed on March 22, 2012).
- Rosen, Stephen Peter, , “Net Assessment as an Analytical Concept,” in Andrew W. Marshall, J. J. Martin, and Henry S. Rowen, eds., *On Not Confusing Ourselves: Essays on National Security Strategy in Honor of Albert and Roberta Wohlstetter* (Boulder, CO: Westview Press, 1991), pp. 283-301.
- Rosen, Stephen Peter, *War and Human Nature* (Princeton, NJ: Princeton University Press, 2005).
- Rosenberg, Alexander, *Philosophy of Social Science* (Boulder, CO: Westview Press, 2007).
- Rougier, Bernard, *Everyday Jihad* (Cambridge, MA: Harvard University Press, 2007).
- Rovner, Joshua, *Fixing the Facts: National Security and the Politics of Intelligence* (Ithaca, NY: Cornell University Press, 2011).
- Russell, Richard L., “Achieving All-Source Fusing in the Intelligence Community,” in Loch K. Johnson, ed., *Handbook of Intelligence Studies* (New York, NY: Taylor and Francis, 2009), pp. 189-198.
- Sabin, Philip, *Lost Battles: Reconstructing the Great Clashes of the Ancient World* (New York, NY: Continuum, 2009).
- Sabin, Philip, *Simulating War: Studying Conflict Through Simulation Games* (New York, NY: Continuum, 2012).
- Sageman, Marc, *Understanding Terror Networks* (Philadelphia, PA: University of Pennsylvania Press, 2004).
- Salamon, Thomas, *Design of Agent-Based Models* (Czech Republic, Tomas Bruckner, 2011).
- Saunders-Newton, Desmond, “Computational Social Science, Operations Research, and Effects-Based Operations: The Challenge of Inferring ‘Effects’ from Dynamic Socio-Physical Systems,” *Analyzing Effects-Based Operations*, Military Operations Research Society, McLean, VA, January 28, 2002.
- Saunders-Newton, Desmond, and Aaron B. Frank, “Effects-Based Operations: Building the Analytic Tools,” *Defense Horizons* No. 19 (October 2002).

- Sawyer, R. Keith, *Social Emergence: Societies as Complex Systems* (New York, NY: Cambridge University Press, 2005).
- Sayer, Andrew, *Method in Social Science* (New York, NY: Routledge, 1992).
- Schelling, Thomas C., *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1980).
- Schelling, Thomas C., *Micromotives and Macrobehavior*, (New York, NY: W.W. Norton and Company, 2006).
- Schrodinger, Erwin, *What is Life?* (New York, NY: Cambridge University Press, 2004).
- Schwartz, Peter, *The Art of the Long View: Planning for the Future in an Uncertain World* (New York, NY: Doubleday, 1996).
- Shachtman, Noah, "Pentagon's Prediction Software Didn't Spot Egypt Unrest," *Danger Room, Wired* (February 11, 2011), <http://www.wired.com/dangerroom/2011/02/pentagon-predict-egypt-unrest/> (accessed on August 6, 2011).
- Simon, Herbert A., *Models of My Life* (Cambridge, MA: MIT Press, 1996).
- Simon, Herbert A., *The Sciences of the Artificial* (Cambridge, MA: MIT Press, 1996).
- Simon, Herbert A., *Administrative Behavior* (New York, NY: The Free Press, 1997).
- Simon, Herbert A., "Bounded Rationality in Social Science: Today and Tomorrow," *Mind and Society*, Vol. 1, No. 1 (2000), pp. 25-39.
- Sims, Jennifer E., "Decision Advantage and the Nature of Intelligence Analysis," in Loch K. Johnson, ed., *The Oxford Handbook of National Security Intelligence* (New York, NY: Oxford University Press, 2010), pp. 389-403.
- Sims, Jennifer E., and Burton Gerber, eds., *Transforming U.S. Intelligence* (Washington, DC: Georgetown University Press, 2005).
- Singh, Abhinav, Dmitri Vainchtein and Howard Weiss, *Limit Sets for Natural Extensions of Schelling's Segregation Model*, June 15 2010, <http://people.math.gatech.edu/~weiss/pub/Sch2June15.pdf> (accessed on November 8, 2012).
- Smith, Timothy J., "Predictive Warning: Teams, Networks, and Scientific Method," in Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 266-280.

- Squazzoni, Flaminio, *Agent-Based Computational Sociology* (West Sussex, United Kingdom: James Wiley and Sons, 2012).
- Stalnaker, Robert, *Mere Possibilities: Metaphysical Foundations of Modal Semantics* (Princeton, NJ: Princeton University Press, 2011).
- Steinberg, James B., “The Policymaker’s Perspective: Transparency and Partnership,” in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 82-90.
- Steury, Donald P., ed., *Sherman Kent and the Board of National Estimates: Collected Papers* (Washington, DC: Central Intelligence Agency, 1994).
- Stogdill, Ralph M., ed., *The Process of Modeling-Building in the Social Sciences* (New York, NY: W.W. Norton and Company, 1970).
- Straus, Susan G., Andrew M. Parker, James B. Bruce and Jacob W. Dembosky, *The Group Matters: A Review of the Effects of Group Interaction on Processes and Outcomes in Analytic Teams* (Santa Monica, CA: RAND, April 2009).
- Strauss, Walter J., “The Nature and Validity of Operations Research Studies, With Emphasis on Force Composition,” *Operations Research*, Vol. 8, No. 5 (September/October 1960), pp. 675-693.
- Surowiecki, James, *The Wisdom of Crowds* (New York, NY: Doubleday, 2004).
- Taber, Charles S., and Richard J. Timpone, *Computational Modeling* (Thousand Oaks, CA: SAGE Publications, 1996).
- Telhami, Shibley, and Michael Barnett, eds., *Identity and Foreign Policy in the Middle East* (New York, NY: Cornell University Press, 2002).
- Tetlock, Philip E., and Aaron Belkin, “Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives,” in Philip E. Tetlock and Aaron Belkin, eds., *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives* (Princeton, NJ: Princeton University Press, 1996), pp. 3-38.
- Tetlock, Philip E., and Aaron Belkin, eds., *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives* (Princeton, NJ: Princeton University Press, 1996).
- Tetlock, Philip E., Richard Ned Lebow, and Geoffrey Parker, “Unmaking the Middle Kingdom,” in Philip E. Tetlock, Richard Ned Lebow, and Geoffrey Parker, eds.,

- Unmaking the West: "What-If?" Scenarios That Rewrite History* (Ann Arbor, MI: University of Michigan Press, 2006), pp. 1-13.
- Tetlock, Philip E., Richard Ned Lebow, and Geoffrey Parker, eds., *Unmaking the West: "What-If?" Scenarios That Rewrite History* (Ann Arbor, MI: University of Michigan Press, 2006).
- Thagard, Paul R., "Why Astrology is a Pseudoscience," in E. D. Klemke, Robert Hollinger, and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), pp. 66-75.
- Thompson, John N., "Coevolution," in Simon A. Levin, ed., *The Princeton Guide to Ecology* (Princeton, NJ: Princeton University Press, 2009), pp. 247-252.
- Tilly, Charles, *Coercion, Capital and European States: AD 990-1992* (Malden, MA: Blackwell Publishing, 1992).
- Transformation Study Group, *Transformation Study Report: Executive Summary* (Washington, DC: Department of Defense, April 27, 2001).
- Transformation Study Group, *Transformation Study Report: Transforming Military Operational Capabilities* (Washington, DC: Department of Defense, April 27, 2001).
- Treverton, Gregory F., *Reshaping Intelligence for an Age of Information* (New York, NY: Cambridge University Press, 2003).
- Treverton, Gregory F., "Intelligence Analysis: Between "Politicization" and Irrelevance," in Roger Z. George and James B. Bruce, eds. *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, DC: Georgetown University Press, 2008), pp. 91-104.
- Trichet, Jean-Claude, *Reflections on the Nature of Monetary Policy Non-Standard Measures and Finance Theory*, November 19, 2010, <http://www.ecb.int/press/key/date/2010/html/sp101118.en.html> (accessed on October 12, 2012).
- U.S. Senate, Senate Select Committee on Intelligence, Report on the U.S. Intelligence Community's Prewar Intelligence Assessments on Iraq, (Washington, DC: United States Senate, 7 July 2004).
- United States Institute for Peace, *Strategic Economic Needs and Security Exercise (SENSE)*, <http://www.usip.org/sense> (accessed on September 10, 2012).

- Van Der Heijden, Kees, *Scenarios: The Art of Strategic Conversation* (New York, NY: John Wiley and Sons, Ltd., 2009).
- von Clausewitz, Carl, *On War* (Princeton, NJ: Princeton University Press, 1984), Michael Howard and Peter Paret, trans.
- Wack, Pierre, "Scenarios: Uncharted Waters Ahead," *Harvard Business Review*, No. 85516, (September-October 1985), pp. 73-89.
- Walt, Stephen M., "The Renaissance of Security Studies," *International Studies Quarterly*, Vol. 35, No. 2 (June 1991), pp. 211-239.
- Walton, Timothy, *Challenges in Intelligence Analysis: Lessons from 1300 BCE to the Present* (New York, NY: Cambridge University Press, 2010).
- Waltz, Kenneth N., *Theory of International Politics* (Reading, MA: McGraw-Hill, 1979).
- Ward, Michael Don, ed., *Theories, Models, and Simulations in International Relations: Essays in Honor of Harold Guetzkow* (Boulder, CO: Westview Press, 1985).
- Weinberger, David, *Too Big to Know: Rethinking Knowledge Now That the Facts Aren't the Facts, Experts are Everywhere, and the Smartest Person in the Room in the Room is the Room* (New York, NY: Basic Books, 2012).
- Weiss, Gerhard, ed., *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence* (Cambridge, MA: MIT Press, 1999).
- Wendt, Alexander, *Social Theory of International Politics* (New York, NY: Cambridge University Press, 1999).
- Wikipedia "MONIAC Computer," http://en.wikipedia.org/wiki/MONIAC_Computer (accessed August 5, 2011).
- Wilensky, Uri, *NetLogo*, Center for Connected Learning and Computer-Based Modeling (Evanston, IL: Northwestern University, 1999), <http://ccl.northwestern.edu/netlogo/> (accessed on December 5, 2012)
- Wilson, Andrew, *The Bomb and the Computer: A Critical History of Wargames* (New York, NY: Delacorte Press, 1969).
- Wirtz, James J., "Intelligence to Please? The Order of Battle Controversy During the Vietnam War," *Political Science Quarterly*, Vol. 106, No. 2 (Summer 1991), pp. 239-263.

- Wirtz, James J., "The American Approach to Intelligence Studies," in Lock .K. Johnson, ed., *Handbook of Intelligence Studies* (New York, NY: Taylor and Francis, 2009), pp. 28-38.
- Wohlstetter, Roberta, *Pearl Harbor: Warning and Decision* (Stanford, CA: Stanford University Press, 1962).
- Wolfowitz, Paul, "Comments: Paul Wolfowitz," in Roy Godson, Ernest R. May and Gary Schmitt, eds., *U.S. Intelligence at the Crossroads: Agendas for Reform* (Washington, DC: Brassey's, 1995), pp. 75-80.
- Wolfowitz, Paul, *Deputy Secretary Wolfowitz Interview with Sam Tannenhaus*, *Vanity Fair*, May 9, 2003, <http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2594> (accessed on May 30, 2012).
- Wolfram, Stephen, *Cellular Automata and Complexity* (Boulder, CO: Westview Press, 1994).
- Wooldridge, Michael, "Intelligent Agents," in Gerhard Weiss, ed., *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence* (Cambridge, MA: MIT Press, 1999), pp. 27-77.
- Ziman, John, "What is Science?" in E. D. Klemke, Robert Hollinger and David Wyss Rudge, with A. David Kline, eds., *Introductory Readings in the Philosophy of Science* (Amherst, NY: Prometheus Books, 1998), pp. 48-53.
- Ziman, John, "Getting Scientists to Think About What They are Doing," *Science and Engineering Ethics*, Vol. 7, No. 2 (2001), pp. 165-176.

Curriculum Vitae

Aaron B. Frank holds a BA in Political Science from Boston University and an MA in Security Studies from Georgetown University. He also attended the Complex Systems Summer School at the Santa Fe Institute and participated in numerous academic seminars, conferences and short courses.

In addition to his academic studies, Mr. Frank has more than fifteen years' experience working in the national security community and worked at Booz Allen Hamilton, BAE Systems, LMI and George Mason University.

Mr. Frank lives in Fairfax, Virginia with his wife Amy, son Alex and dog Jackie-O.