SANDIA REPORT
SAND2006-0860P
Unlimited Release
February 2006

An Agent-Based Model Component to a Framework for the Analysis of Terrorist-Group Dynamics

George A. Backus and Robert J. Glass

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2006-0860P
Unlimited Release
Printed February 2006

An Agent-Based Model Component to a Framework for the Analysis of Terrorist-Group Dynamics

George Backus (gabacku@sandia.gov)
Robert J. Glass (rjglass@sandia.gov)

Sandia National Laboratories
PO Box 5800, Albuquerque, New Mexico 87185

January 23, 2005

Work Performed for the Department of Homeland Security
Motivation & Intent Thrust Area

Abstract

This document presents the conceptualization of an agent-based, simulation framework that allows the use and testing of various social and behavioral science approaches for understanding the motivation and intent associated with terrorist activities. The framework design provides a LEGO™ -style toolbox that can convert sophisticated SME theses on individual and social behavior into computationally tractable, mathematical representations. Through parameterization, the reconfigurable framework can then simulate the dynamics of any particular group or interacting collection of terrorist groups.

Table of Contents

1.0 Introduction

This document describes an agent-based approach for analyzing terrorist group dynamics. It starts with an overview of issues associated with modeling social behavioral phenomena, followed by a general (non-mathematical) outline for a "building block" methodology that can translate the language of sociological, political, and psychological research on terrorism and related efforts into a computationally tractable representation. Appendices furnish more in depth, and often mathematical, discussions on the building blocks and their supporting research.

2.0 Background

The Department of Homeland Security, through its Motivation & Intent (M&I) Thrust Area, intends to use computerized models to ultimately improve the efficiency of intelligence analysts as they attempt to assess terrorist threats. A model is a machine, and like a machine, it can only perform the function for which it was built if its design and construction were sufficient. The successful development of an M&I model requires a clear understanding of how to combine modeling methods with the subject matter of interest.

2.1 Mental Versus Mathematical Models

All human motivation, intent, and action is based on the use of models. For the most part, they are the consequence of mental models routinely expressed in spoken language, but often they are never explicitly verbalized (Doyle 2001). Mathematics is a language much like any other. A simulation model is the translation of a mental model into a useful, mathematical formalization (Sterman 1988). Unlike many spoken languages, the use of the language of mathematics in a model allows paradox, but not inconsistency. Models are therefore useful to validate analysis approaches and to determine the implications and consequences of the approaches (Sterman 2000, Chapter 1).

Terrorist behavior is a dynamic process. It is the consequence of both life-long and near-term events. The socioeconomic and geopolitical environment helps shape the perceptions and responses of an individual. In turn, when the responses affect the vested interests (such as those of the government), those interests counter-respond, within their capability, in what can lead to a vicious cycle (Silke 2003, p51). Essentially all real world systems are dominated by feedback, where interacting entities affect each other and drive future responses. The human mind is ill-adapted to understand the evolution and ramifications of these feedback dynamics (Sterman 1989). The only means to enhance our mental models and thoroughly comprehend the dynamic repercussions is through the use of simulation models (Sterman 2000, p29). Modern simulation methods allow the full use of qualitative perspectives (such as culture) and quantitative information (such as violent-group activities) to address pressing social-behavioral problems and augment human cognitive limitations (Bednar and Page, 2005).

"The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real word." (Simon[1] 1957, p198)

"Simulation is the only practical way to test these [mental] models. The complexity of the mental models vastly exceeds our capacity to understand their implications." (Sterman 2000, p38).

The conundrum of terrorism leads to its categorization as a wicked problem (Rittel 1973, Richey 2005).

"A wicked problem usually has the characteristics that many people have tried and failed to solve the problem, that most attempts to solve the problem actually make the problem worse, and that nobody really agrees on exactly what the problem actually is! Wicked problems typically arise in feedback systems, and almost any interesting system is a feedback system." (Palfreyman 2003)

Addressing wicked problems requires a recognition that "priors" (preconceptions based on assumption or intuition) can distort the ability to see or accept alternative formulations of, or approaches to, the problem (Meadows 1980). New methods in Bayesian analysis (see Appendix 1) appear to allow the determination of what causal mechanisms a simulation should include or, conversely, cannot reject in a social-science model (McKim 1997, Spirtes 1993, Glymour 1999). A modern statistical method called cointegration (see Appendix 2), further allows the determination of those feedback processes that dominate behavior (Engel[2] 1991; Hendry 1993,1995). Sensitivity analysis methods have been successfully applied to combined societal-behavioral and physical systems that are dominated by feedback (Backus 1985). Empirically, the feedback assures that only a few mechanisms control behavior under any given condition and these key mechanisms are readily determined (McKay 1976). Analysts can than concentrate on those key elements. None of the above noted methods have yet become widely used within the social-psychological (behavioral) community. As such, conclusions drawn about what is ***not*** a cause of terrorism, using current statistical practices, might be misleading (Gigerenzer 1993).

2.2 Agents and Social Movement Theory

Several evaluations of terrorist activities include consideration of alternative explanations for terrorist behavior (McCauley 2002, 2006; Borum 2004). These evaluations make it clear that group dynamics act to magnify the propensities of individual behavior that can lead to terrorist actions. The apparent complexity of terrorist behavior has lead to an

[1] Awarded Nobel Prize in 1978
[2] Awarded Nobel Prize in 2003

emphasis on group dynamics, as embodied in social movement theories (Beuchler 1995). From a modeling perspective, social movement theory acts largely as a classification process that offers descriptive capabilities delineating characteristics to which the collective response of simulated agents (individuals) should correspond. That is, group dynamics come from the interacting individuals forming the group.

Agent-based modeling simulates the changing interactions among individuals and their decisions. Agent-based modeling has the ability to produce the collective (group) emergent behaviors from evolving individual (agent) behaviors. Thus, agent-based modeling is particularly suited for the analysis of terrorist group dynamics. Within an agent-based simulation, simple individual mechanisms, such those associated with relative-deprivation or frustration, that alone cannot account for the group level action, can interact with other primal behaviors to produce the group behaviors that social movement theory attempts to describe. Agents behave differently in different environments and under the influence of different histories (Bednar and Page 2005). The varied and seemingly inconsistent histories of various terrorist groups could simply be a consequence of the specific history and environment that shapes the collective agent trajectory. A few simple mechanisms may explain a wide variety of dynamics, and can thereby offer an understanding (and control) of terrorist motivation and intent. Additionally, violence comes from a relatively small element of any society. Agent-based models simulate a distribution of individuals. In essence, the agents reflect the full spectrum of possibilities and can produce the full spectrum of actualizations.

While groups can be represented as an interacting and evolving set of individuals, individuals (agents) can be represented by a primal set of behavioral mechanisms. The mechanisms are the language that expresses the individual behavior and its connection with other individuals (and the environment). That language must be mappable to/from the language that subject matter experts (SMEs) use to describe behaviors and dynamics. For human behavior to be so complex, it must be composed of a relatively small number of simple mechanisms. If there were a large number of complicated mechanisms, the mechanisms would only work in the situation for which they were designed. Simple mechanisms can fit together and interact in near infinite variety. The concept is no different than that of jig-saw puzzle, where complex pieces can only make one picture, but simple pieces such as colored triangles, can make any picture desirable (albeit, with a few negligible, rough edges). Lego™ pieces can portray very complex shapes because they allow both two and three dimensional constructs. (They even allow four dimensional concepts if the pieces are hinged, and, thus, can vary over time). The Lego™ metaphor seems appropriate for an agent-based framework. The use of a few key (behavioral) building blocks can decompose the logic of a SME. The replicated use of those blocks can then form the representation for any group or interacting set of groups. If the building blocks are well designed, their parameterization with historical data will result in a realistic and useful representation of the group relative to the future motivation and intent (M&I) dynamics of interest.

Formal methods exist to ensure the agent-based modeling framework captures both the social-science basis for analysis as specified by alternative SME perspectives, as well as a useful description of arbitrarily-specified focus groups. David Hendry developed the concept of encompassing (Hendry 1993, Chapter 13). "A model 'M' is said to encompass another model 'N' if the former can explain the results obtained by the latter" (Florens and Hendry 1996). The "models" can be mental or computerized, as long as there are data to distinguish variance among the alternatives. The model design used in this work allows the reconfiguration of model agents to conform to a SME perspective and then the parameterization to conform to the focus groups. It encompasses the universe of agent-based terrorism simulation.

2.3 Modeling a Problem

A model is not a model of system, but rather the model of a problem. Once a problem is determined, the finite selection of salient features the model must contain becomes manageable. Without a "problem," a model design has no focus or end point. Most of all, without a problem focus, a model has no definition of "useful" (Sterman 2002). Different paradigms (modeling methods) allow different views of the problem (Meadows 1980). Despite the fact different paradigms offer "different" solutions, they can be complimentary. For example, Bayesian Prospect Analyses and Monte-Carlo Markov Chain analysis provide the conditional probabilities for parameterizing the endogenous, dynamic, conditional probabilities within an agent–based model. Optimization and game-theoretic analysis determine the idealized goals that intervention scenarios within an agent-based model should attempt to realize.

Although not part of the current focus, behavioral agent models with feedback can additionally simulate mid- and long-term dynamics that allow control over the formation of M&I within the target population (focus groups). They can find leverage points affecting M&I formation or change in M&I direction, as well as determine time constants and constraining feedback on societal/behavioral drivers. Given uncertainty and the limits of external control (intervention) functions, they define robust strategies that produce stabilizing dynamics. With sensitivity analysis, agent-based models, can recognize how operational and tactical responses to M&I might affect future M&I conditions/responses within the target population.

The immediate function of the modeling framework is to determine the M&I of specific focus groups. The models are calibrated to reflect the historical dynamics of specific groups with a perspective on future activities. While historical reproduction facilitates the initial parameterization of a model, it provides no measurable metric for having confidence in the model (Sterman 2000, Chapter 21). Only formalized testing and extended review improves the confidence in a model's use for addressing a ***specific*** problem/question. The evaluation metric of the framework currently corresponds to the faithfulness in its inclusion of SME perspectives and the reasonable resemblance of its results to the focus group behaviors. As such, the model requires an apple-to-apple comparison of SME perspectives in a defined format with uniform definitions. Similarly, the models require historical time-series information about each group in a defined format with uniform definitions. Only an apple-to-apple comparison avoids the ambiguities that render a model meaningless.

2.4 Generalized Agents

Figure 1 illustrates a generalized agent. An actual computerized agent may only have a few of the components or may have several components of the same type. Like a human (or any organism), the agent receives information from its environment. The environment may be another agent or the general physical or governmental conditions. Given filters and distortions, it assimilates the conditions, based on its understanding of the context of the information. These perceived conditions add to the experience of the agent and may cause changes in the agent that affect future cognition. Given a set of available alternative actions, the agent considers the perceived utility of each action and agent decides a behavior (including no action). Any action engages the environments and may cause the environment to change or counter-respond. The change in the environment (as caused by the agent) then feeds back on the agent.

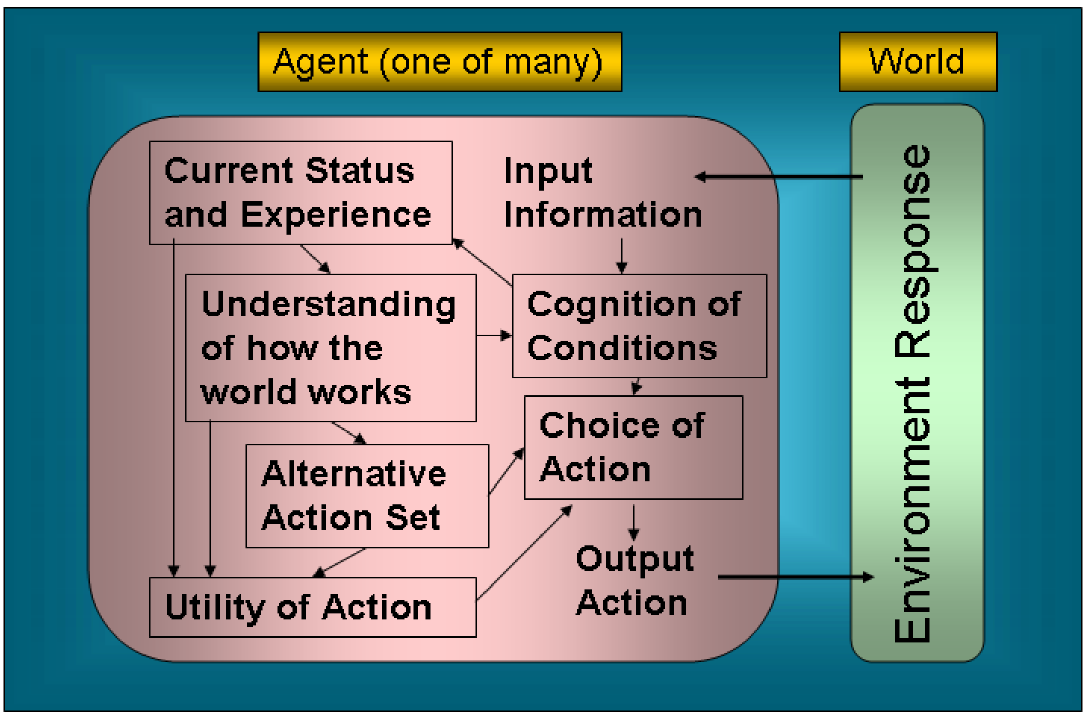


Figure 1: A Generalized Agent

Numerous packages allow the specification and simulation with agents, such as RePast and AnyLogic. When the relationship of agents within a network becomes an important consideration of the problem and dynamics, variants on these tools, such as Loki for RePast, explicitly capture the network dynamics.

3.0 Agent-Based Behavioral Analysis

In the domain of terrorism (and most all other behavioral spheres), an agent compares a goal to the environment. If there is an unacceptable difference between the goal and reality, the agent makes a decision that is often associated with an action that changes the environment. As depicted in Figure 2, there is a feedback loop (often with delays) between the action and the environment.

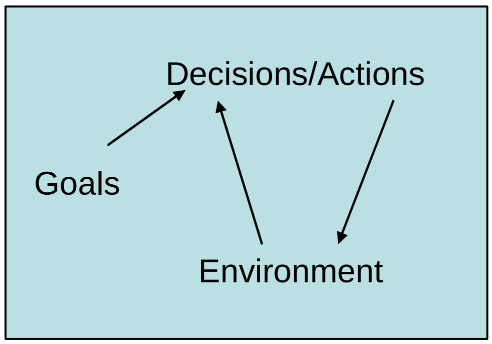


Figure 2: The basic environment-goal-decision relationship

Figure 3, shows a more comprehensive view. The goal itself is a function of the experiences the environment produces. There can be additional side affects. And the environment also changes as a multitude of agents interact.

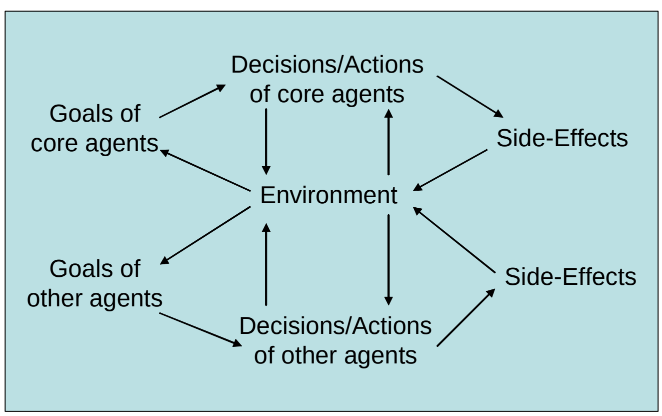


Figure 3: The agent-environment-agent relationship

The models here will look at the dynamics of terrorist groups to determine the motivation and intent toward violent activities. The design discussed below, 1) determines the causal drivers of such motivation and intent, 2) determines the probability of political, violent, or passive activities, 3) terrorist group formation and capability development, 4) the escalating terrorist and governmental response, 5) the varied impacts of media, religion, the internet, culture and other "technologies, and 5) the influence of socioeconomic and geopolitical change on terrorist group sustainability. The model can address all these issues over the extended "life' of a group, including fragmentation and interactions with other groups. The current model design does not include the physical consequences of violent actions. Rather, it only addresses those aspects that further or reduce the motivation and intent itself. Individuals in the model are representative. The simulated groups can have a one-to-one correspondence with the actual group and government/nation simulated. Physical drivers to the system, such as economic growth, foreign interventions, and the unpredictable action of a single momentous leader are also exogenous to the model.

As will be discussed in detail below, the current model design only includes four primal building blocks (components): expectation-formation, response-behavior (choice), coping-capacity, and opinion adjustment. Opinion-adjustment combines coping-skill and behavioral-response dynamics into a single component applicable to a key subset of terrorist dynamics. With this small set of building blocks, it appears the model can capture all the causal dynamics noted in the literature, as referenced in previous section. The components are a-cultural. That is, the mechanisms represent primal behaviors common to all humans, upon which parameterization imposes cultural characterization, as appropriate.

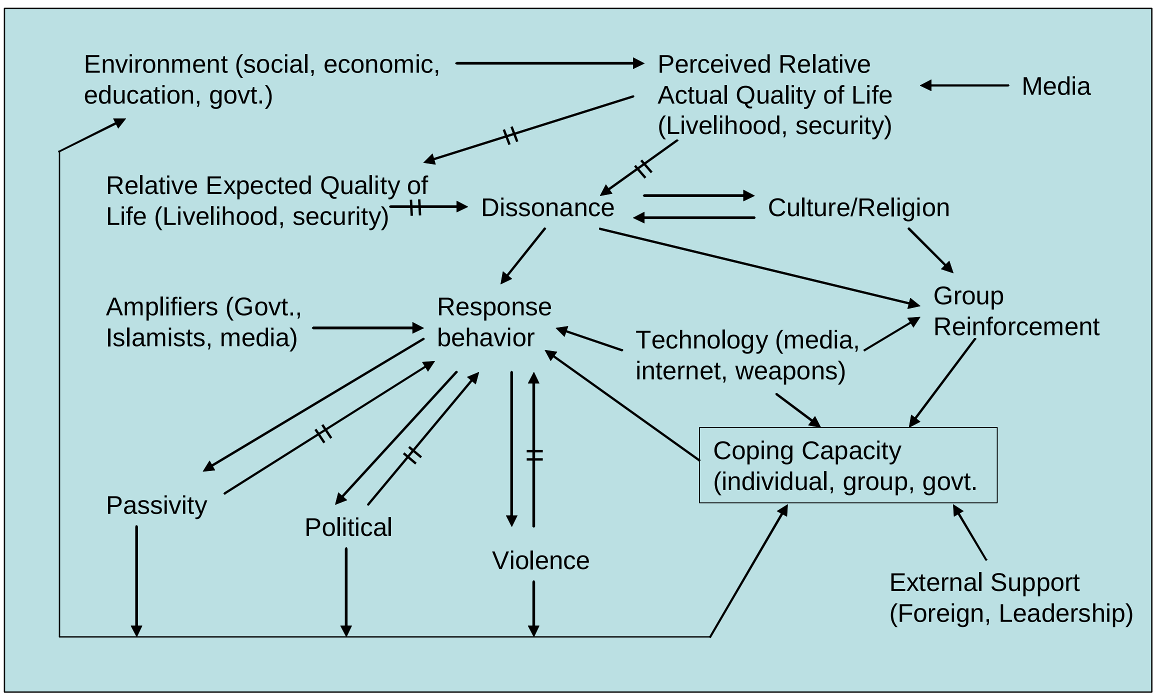


Figure 4: Model Components.

Figure 4 presents an overview of an illustrative model structure that contains examples of the building block utilization and captures many generally accepted social-behavioral phenomena. Each item in Figure 4 can be associated with an agent. Each agent can be an individual or an institution (i.e., a government). Starting in the upper left-hand corner of the diagram, environmental conditions such as employment (income), education, social conditions, and government control determine current realizations and expectations of security and livelihood. Security and livelihood are used as an aggregate metric to capture the key aspects of what might be considered Maslow's Taxonomy (Maslow 1962).

The difference between expectations and perceived reality produces a dissonance.[3] The cross-line in the arrows of causality in Figure 4 indicates delays in response. It takes time to gestate and incorporate information. Religion and culture present a form of coping skills. As dissonance occurs, people seek (or withdraw to) religious or cultural traditions that allow a rationalization of conditions. These psychological anchor points act to mitigate the dissonance. The simple representation noted in Figure 4 denotes the choice to seek religious and cultural moral support. Secondary aspects of religions will become clear momentarily. Depending on 1) the ability to act (coping-capacity), 2) the existing effectiveness of violence or political activities, 3) outside influences, and 4) the level of dissonance, the individual makes choices toward violence, passivity, or political activity. Group dynamics can dominate the response-behavior. The group amplifies the coping capacity of the individual and reinforces a predilection toward continued association with the group and its function. In this example, the groups represent proponents of the three responses simulated (passivity, political, violence). (The model can include additional responses and distinctions as actual or analogous data allow.)

External (exogenous) support from foreign governments/resources and exceptional leadership personalities can dramatically augment the amplifying affect a group has on an individual. The internet makes it easier to join a group both in effort and risk terms. (In economics, this phenomenon is called reducing the transaction and hurdle costs). The media can distort the information an individual uses to make a choice. Weapon technology also amplifies the coping capacity of an individual, just as a power tool amplifies the capability of its user. It is not only the violence or political activism that an individual sees in his social environment, but also the counter (or precipitating) violence or political activity of the government (or government surrogate) that affects the choices made. The reaching out to religious or cultural organizations that resonate with the level of dissonance act primarily to increase the probability of contact with groups that reinforce rather than mitigate the dissonance. Again, the Internet and media also increase the probability of contact of similarly dissonant souls. The added affect of Islamist education will be discussed later.

From a sociological and psychological level, dissonance is the measure of motivation and the behavioral response is the measure of intent. From a DHS perspective, the intent is multivariate, and the coping capacity and group-reinforcement qualify as operational measures of motivation. The sections that follow expand on the logic noted above and implied in Figure 4.

3.1 Expectations

The psychologist Hogarth (1987) notes that almost all decisions are based on expectations. Further, the shortcomings of human cognition and information assimilation mean that decisions are based on limited information and selective perceptions (Miller

[3] The often used term "grievance" is a more specific (negative only) aspect of dissonance.

1956). Additionally, new work shows that humans use information to build up "priors" that are then used as a reference bias for subsequent decisions under the normally existing conditions of uncertain and deficient data (Griffiths and Tenenbaum 2006).

Business planning explicitly recognizes expectation formation (Backus 2001). Simulations using adaptive expectation formation seem to closely match actual expectations (Sterman 1988, Naill and Backus 1977, Backus 1977). Appendix 4 describes the equations used to produce expectations. These equations are based on exponential smoothing of information to filter out the noise and to find the trend. Exponential smoothing is consistent with the Erlang distribution common to many everyday events. The human mind appears to have evolved to explicitly deal with such distributions (Griffiths and Tenenbaum 2006).

Humans form expectations and make decisions about issues that are worth the effort. In the current conception of the model, separate considerations of security and livelihood act as surrogates for the key issues driving individual and societal evolution. Having adequate food, housing, income, and employment (livelihood) means little if you won't live long enough to enjoy it (security). An education brings with it expectations of a better livelihood. Lawlessness and excess government repression add to a sense of insecurity. Weighted sums of socioeconomic and security indicators produce livelihood and security indices, respectively. A multiplicative combination of security and livelihood indices produce a quality of life (QOL) index that captures the key dynamics characteristics important to decision making (Keeney 1976, Meadows 1974). Over time, the perceived current QOL evolves into the expected QOL.

The media can also filter or distort information. In the expectation formation part of behavior, media information exaggerates or understates what would be the normal cognition of information identifying a condition (Atran 2006).

Human expectations are also affected by the changing conditions of the humans around them. The opening of a new industry brings expectations of employment and income, but if those around you receive a better job than you do, there can be a large gap between your perceived condition and your expectations -- that would be far different than those from you getting a poor job, but those around you receiving an even lesser position. The expectations are relative, and usually compare your perceptions to the larger societal fabric around you. In this perspective, deprivation is nothing in isolation. It only takes on behavioral meaning when the relative perception of current conditions are at significant odds with expectations. In this context, even well educated, upper income males may consider themselves victims of relative deprivation (Gurr 1970) and take drastic action in response.

3.2 Dissonance

The dissonance between perceived conditions and expectations produce alienation. This alienation can lead to antisocial behavior (Sageman 2004). Within the model, dissonance is strictly defined as the proportional difference between perceived conditions and expected conditions. Only when there is a gap between perceived reality and expectations, is there an incentive to act. The recent, yet obviously inevitable, social unrest among Chinese workers (Associated Press 2006), is not the result of poverty or persecution, per se, but rather the change in expectations relative to perceived conditions. Many authors indicate that dissonance associated with social and political expectation acts as the starting point on the path to terrorism (Crenshaw 1995 – Introduction; Drummond 2002; Silke 2003, p98)

Religion and culture affect dissonance. Embracing either can improve apparent coping capacity (discussed later). The choice to seek religious and cultural-support increases with dissonance (Dennet 2006). The direct impact is to reduce dissonance and temper the actions that dissonance might engender. As will be noted later, large negative indirect impacts are also possible. This "religion" dynamic is included for completeness and to illuminate the multiple impacts such phenomena as religion and media play in terrorist dynamics. As such, only a simplified coping-capacity plus response-behavior representation of religion is included in this part of the model. Because expectation formation is a filtering (long-term) process, even after conditions change, dissonance changes slowly – as experienced in post-war Iraq.

3.2 Behavioral Response

There is a large and robust body of literature on how to model human choice. One such approach, Qualitative Choice Theory (QCT), has recently garnered the Nobel Prize and the method remains valid whether working with groups or an individual, or rational or irrational behavior (McFadden[4] 1982). Humans make choices among alternative (DeNardo 1985). They use preferences and non-rational considerations, along with the results of previous decisions feedback, to affect future decisions (McCormick 2003, Kahneman[5] 2003a). Humans use a limited amount of data with limited validity by applying heuristic rules (Tversky and Kahnaman 1974, Kahneman 2003b). The mathematics of QCT robustly accommodates the full spectrum of preferences and choice-considerations with minimal data requirements. Most importantly, it accurately portrays extreme decisions made under extreme conditions – as well as trivial, routine decisions (See Appendix 4 for an extended discussion of QCT).

QCT provide a probabilistic assessment of decisions for individuals and a fractional assessment at the group level. The variation in choice comes from uncertainty about the key inputs to decisions. The media can change both the perceived uncertainty and the nature of the actual information itself. Islamic education (e.g., a ranting anti-Western mullah) has the same affect as the media in-kind, but different in degree. With respect to behavioral responses, the internet is also simply a media source. For group-reinforcement, the internet plays a larger role, as noted below.

The violence or political activities of the governments (or other groups) affect the decision to pursue the same. The prior political or violent acts of a group or individual

[4] Awarded Nobel Prize in 2000
[5] Awarded Nobel Prize in 2002

reinforce additional acts of the same nature. It is not much different than buying a new technology (e.g. iPod™) once you know others have done it. If you have success with HP™ computers you will continue to buy them in the future. In the 1920's, if the government oppressively enforced prohibition, it might have kept you away from a speakeasy, but with peer pressure, you might still have gone.

The selection of a choice also depends on the ability to make that choice. It depends on the coping capacity needed to execute the choice. The choice to oppose government/social pressures or to take a great risk requires significant coping capacity. If you feel weak and helpless, you will not oppose the government. If you feel you have an army or your god behind you, no choice seems too extreme. Committing an act makes the next act easier. In an environment of violence, it is hard to break out of a cycle of violence (DeNardo 1985). As will be discussed below, the choice to become part of a group also has its "lock-in" dynamics. Conditions can allow an individual to break away from a group or to cause the group to degenerate, but ceteris paribus, the pressure grows to maintain the group (Olson 1984, 2000). Within the specified QCT logic, both the loop of violence (or politics) and the loop of group reinforcement take time to change – just as evidence dictates (della Porta 1999). The cycle of marginalization forces continued marginalization and alienation.

3.4 Coping Capacity

Anger, or even hate, to the point of wanting to kill somebody, will be void of action if there is no perceived capacity to carry out the act. The psychological coping capacity of individuals determines their ability to respond to conditions. The adaptation to environment comes from changes in coping skills (Helson 1964). If new conditions are excessive compared to coping skills[6] the individual will succumb to the external pressure (flight mode). If the change is within the normal operating range of the coping skills, the individual will counter the new pressure (fight mode). If the new condition is trivial compared to the coping skill, the individual does little and need to do little to (successfully) respond. If the environment provides challenges that are far below the coping skill of an individual, the coping skill atrophies. If the environment produces conditions far in excess of the coping skills to overcome them, there is no ability to improve coping capacity. If the challenge is slightly larger than the average coping capacity, the coping skills increase over time to match the environment. It is easy to make enemies (and drug-resistant diseases) strong by presenting them with escalating, but survivable, challenges – Iran being a current example. The combining of dissonance (the motivation/desire to act) with the limits of coping skills automatically explains many of the seeming paradoxes associated with terrorist histories. Some researchers argue that coping dynamics can help explain all behavior (Moos 2002). It is much as Friedrich Nietzsche stated "Was mich nicht umbringt, macht mich stärker." ("That which does not kill me, makes me stronger.")

In the current model design, there are coping skills of passivity, political response, and violence. The pressure that each of these (government and surrounding social) activities

[6] Coping skills and coping capacity are used interchangeably reflect both its cumulative and dynamic characteristics.

impose consequently results in an individual changing his/her coping capacity. The media, the internet, weapons, and external human support can change the perspective a person on how capable or effective he/she is, and thereby modify the normal evolution of coping capacity. External support can take the form of charismatic leaders, foreign government support, and/or local group support. The existence of external foreign support or an exception leader are exogenous to the model. The group dynamics are, however, endogenous as discussed in the next section.

Mathematically tractable static and dynamic representations of coping skills were finally developed in the 1980s (Lazarus 1984, Backus 1980). Appendix 5 contains a detailed discussion of the coping capacity logic used here to model terrorist motivation and intent.

3.5 Group Reinforcement

Although the choice to be in a group is still a choice (a response-behavior in the language of the model's "building blocks"), it is a choice whose preferences are dominated by evolutionary pressures. Humans are herding mammals (Wilson 1975). From a vulnerable tree dweller, humans evolved coping skills that took advantage of their numbers, along with their co-evolution of communication, by becoming a "social animal" migrating across the savannah (Runciman 2000). Progress in evolutionary psychology indicates that many complex social dynamics (including culture) are outcomes of relatively simple interacting behavioral mechanisms (Barkow 1992, Pinker 2002). Whether in a school of fish, a pack of wolves, or a congregation of people, groups enhance the capability of the individual (Schelling[7] 1960).

The model captures the group dynamics by combining the simple mechanisms noted above – coping capacity and response behavior. The model does not worry about organizational characteristics or unexceptional leaders. Studies have shown that higher level consideration of organizations and leadership usually has minimal dynamic affect on loyalty or activism (Zimbardo 1973). It is the group dynamics and the behavioral reinforcements that seem to be important to terrorist outcomes (McCauley 2002, Sageman 2004).

The model logic implies that individuals generally join group that suit their needs, rather than group coercing allegiance. The group is organic and responsive to the environment. It fulfills a need and evolves accordingly. The individuals in the group reinforce each other. As noted earlier, the coping grows and declines based on the interaction between the individual and the environment (in this instance, the group). Individuals within the group influence and reinforce each other. If there is to be a "growth" toward violence it ccur over time, as several authors have previously noted (della Porta 1991, Sageman 2004).

Figure 5 shows the detailed logic of Group Reinforcement, as shown in aggregate within Figure 4. The group reinforces the coping capacity of the individual toward performing violence or other less detrimental activities. A given individual, who perceives the

[7] Awarded Nobel Prize in 2005

environment such that violence seems the preferred option, will engage a group if 1) he/she comes into contact with the group, and 2) the group has views reasonably similar to his/her own. The conventional probability of making contact corresponds to the fraction of the population an individual meets that are members of the group. The "postulants" can also find each other and form a new group. The internet and media improve the chances of making the contact. As noted earlier, dissonance increases the choice to revert to religion or cultural traditions. (They present a refuge to the familiar that allows one to cope with a the environment by partial escape from it.) By going to a mosque or joining a (ethnic) cultural enclave, the individual self-selects to be in the presence of individuals that have a higher probability of being in a similar condition. This natural self-selection for religion and cultural tradition may cause the illusion that a particular religion or culture has a special association with terrorism. The probability of engagement depends 1) on the drive (dissonance) to join a group, and 2) on the closeness of the group views relative to the view of the individual. (The next section on opinion-adjustment highlights this later phenomenon.)

The internet plays a very important role in reducing the dissonance (energy) to join a group. The risk and effort are minimized and the social "glue" is enhanced because the relationship focuses on the joint attraction with the minimization of negative inter-personal cues. Or as a new study notes: "The internet and email aid users in maintaining their social networks and provide pathways to help when people face big decisions" (Boase 2006). In economics (and the behavior-response formulation) this dynamic corresponds to a minimization of transaction costs.

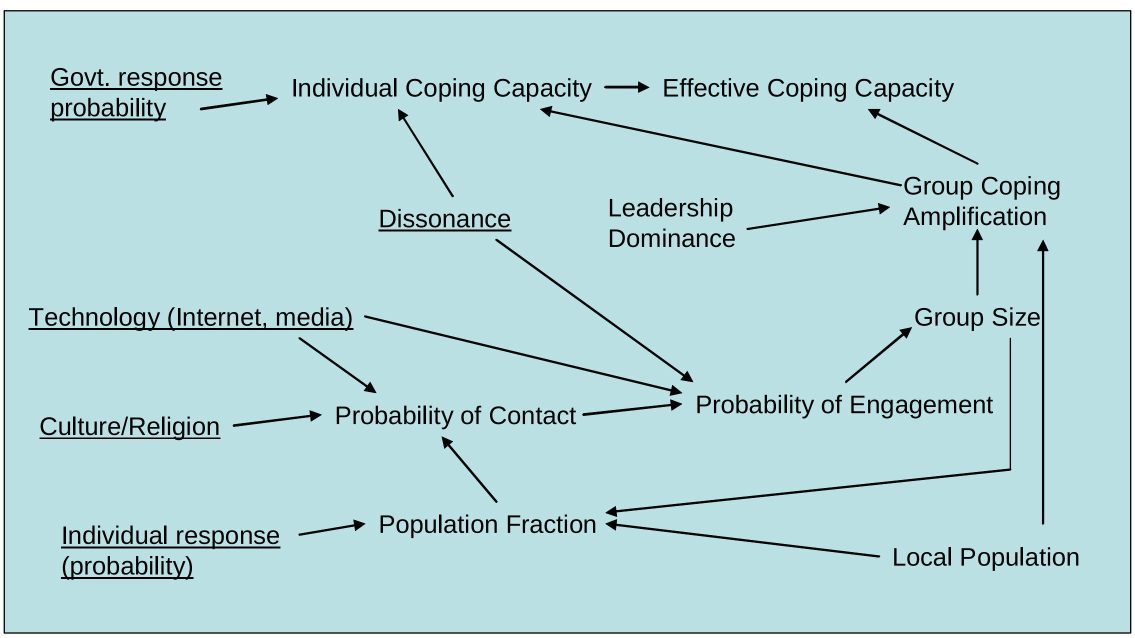


Figure 5: Group Reinforcement Detail

Summing the number of individuals (agents) times the probability of engagement, produces the group size. The backing of the group enhances the coping capacity of the individual (e.g., the stereotypically insecure bully being backed by a group of his "friends"). The individual enhanced by the group can present a much more capable front against government suppression efforts. To the individual, limited government efforts may then help enhance individual coping skills as opposed to overwhelming them. This is why the "group" interaction appears to be the driving force. Group dynamics might instead be the key "amplification" process. (Note that the government is just another agent in the model.)

3.6 Example Implementation of Opinion Adjustment within a Networked Agent-Based Model

This section discusses a simplistic compact representation of some of the above mechanisms and shows that an agent based model that incorporates such a compact representation can produce many of the results expected from more complex descriptions. The model imbeds mechanisms within an explicit social network and is logically extendable to include the full set of mechanisms discussed above. In its final form, it would abstract the essence of self-organized extremist group formation, activation, and dissipation in the context of a structured society (multiply overlapping groups) and where the influence of external and internal drives and perturbations can be evaluated. It would be used to identify system features, attributes, types of perturbations and drivers that result in extremism and determines strategies that can reduce or enhance the likelihood of formation, activation and/or dissipation of extremism.

The simplified model combines the threshold based approach to the agent activation to violence of Epstein (2002) with the dynamics of opinion formation of Weisbuch and coworkers (2002, 2005). Epstein represents an agent's action level by its "Grievance", a combination of two agent state variables: perceived "Hardship" and perceived "Legitimacy" of the local regime or status quo. The term Grievance has an interpretation comparable to that of Dissonance or Motivation. Hardship and Legitimacy are perceptions, or opinions, of individual agents and are the result of interactions with other agents, information from the media, environment, events, etc. In Epstein's analysis, Hardship or Legitimacy are static variables, however the example implementation allows them to be dynamic and to arise through interaction with other agents via the opinion formation process of Weisbuch. Risk of capture or worse, may temper the intent to perpetrate an act of violence. Risk is a function of the individual's inherent aversion to risk and the probability of being discovered and suffering an undesirable outcome. If the motivation, adjusted for risk, exceeds a threshold, an individual will shift to a new (possibly violent) action state. Appendix 6 describes the structure and its logic more fully.

The logic mirrors that of dissonance, but with group formation and reinforcement in a single representation. If an individual comes into contact with another individual and their two views are within a range of each other, then the individuals have enough in common to entertain each other's view. They might compromise on those views, with one becoming more radicalized and one less so. Figure 6 show two individuals with two different opinions (solid colors) and two different thresholds for engagement (shaded area). They will not engage each other and compromise. The axis, for example, could go from “Pro-Western” (0.0) to Violently Anti-Western (1.0).

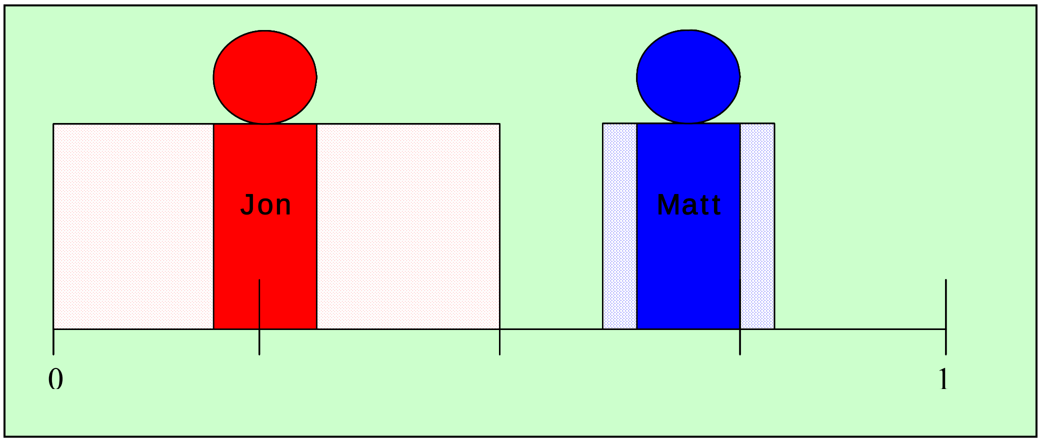


Figure 6: Non-overlapping opinions and thresholds.

Figure 7 shows the situation where and thresholds totally overlap. Plasticity defines how flexible an individual is. If Matt is very inflexible, as implied by the reduced threshold area, it will be Jon who mostly moves his opinion toward Matt’s.

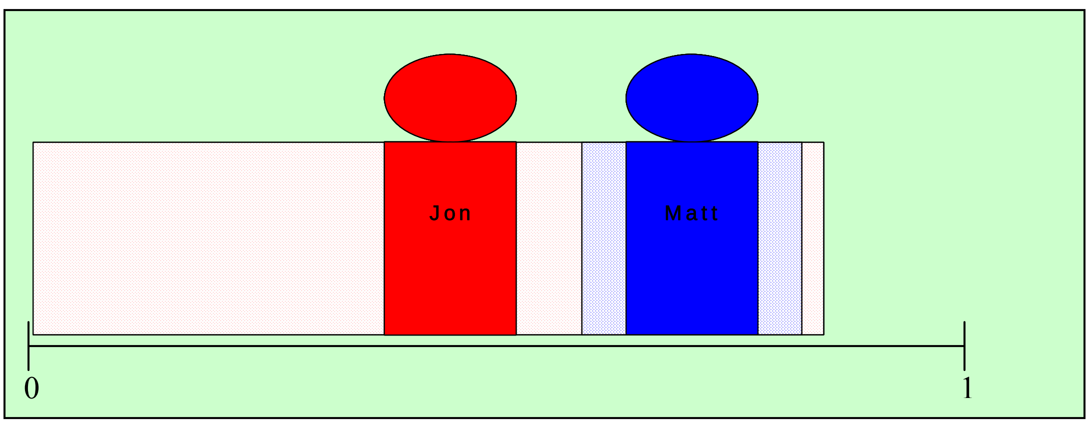


Figure 7: Overlapping Tolerance

If the dynamics only move individuals toward one another, the formation of violent groups only occurs under very specific conditions where the population is initially seeded with a number of low tolerance extremists (Weisbuch et al., 2002, 2005). If, however, a second dynamic of antagonism, where an individual moves apart if the opinions of another individual are too removed from their own, then the rapid formation of extremist groups become possible under more general conditions (Chambers and Glass 2005). The antagonistic response certainly conforms to everyday religious and political debates, and it is consistent with the apparent engagement of “ordinary’ individual into extremists groups (Sageman 2004). It also appears to agree with disconcerting, recent analyses that antagonistic responses are valid for both facts and opinions (Westin 2006).

Figures 8 and 9 show an early model run that includes antagonism, and displays how the legitimacy of varying groups can change, with one group ultimately driving the outcome.

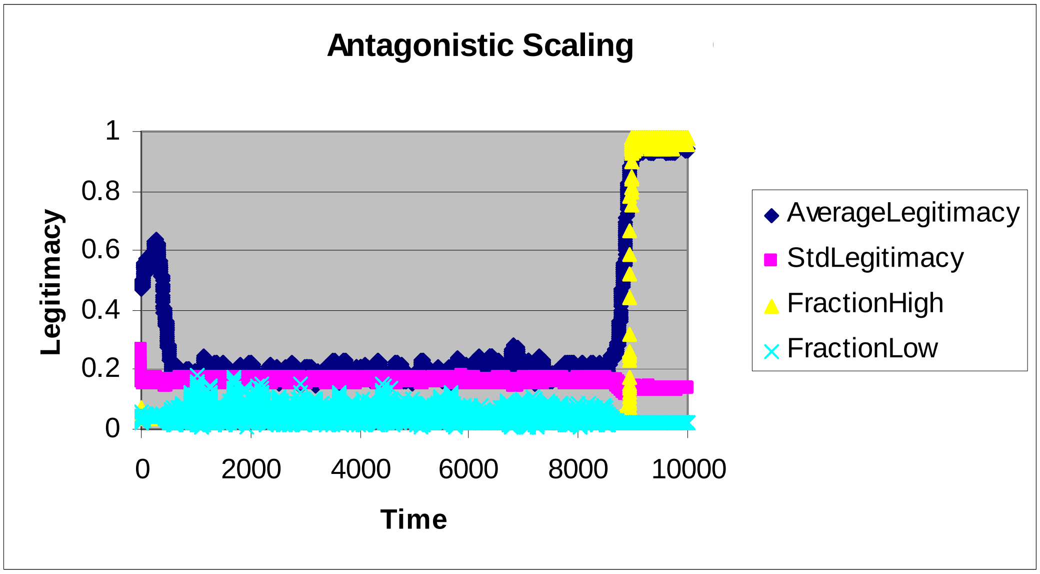


Figure 8: The establishment of a dominant radical group

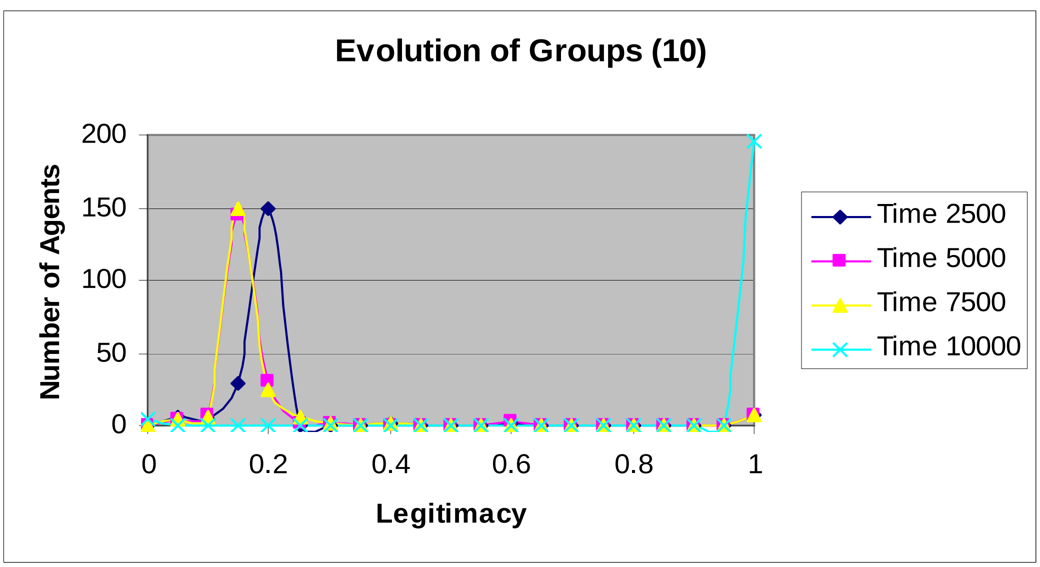


Figure 9: The dynamic transition of group dominance.

The initial modeling utilizes the opinion adjustment logic and is designed to expand to use the other noted mechanisms as data become known and the perspective of subject matter experts are compiled and mapped into mathematical formalizations.

4. 0 Summary

Social behavior is the consequence of individuals acting within a group construct. The dynamics of such processes are inherent in the individuals and the interaction among them and their environment. Individual behavior is, in turn, the consequence of interacting primal mechanisms that evolved to allow human behavior and survival. It appears that simulated agents, composed of the key mechanisms germane to the problem of interest, can imitate the dynamics of actual terrorist-group motivation and intent. A building block approach to these mechanisms allows the translation of subject-matter expertise into the mathematics needed to drive an analysis. This paper addressed recent breakthroughs in the mathematical representation and the understanding of behavior to develop "building blocks" suitable for constructing an agent-based model of terrorist dynamics.

5.0 References

Associated Press (2006) “Chinese Factory Workers, Police Clash” January 23, 2006 http://news.yahoo.com/s/ap/20060123/ap_on_re_as/china_factory_protest

Atran, Scott (2006), “Risk in the Wild: Reassessing Terrorist Threats from the Field, AAAS Annual Meeting, Risk and Society Panel, February 19, 2006, St. Louis, MO

Backus, George A. 1977. FOSSIL1: Documentation. DSD #86. Resource Policy Center. Dartmouth College, Hanover, New Hampshire.

Backus, G. A. (1980). Human Behavior: Culture vs. Biology (Report AFC-1): School of Industrial Engineering, Purdue University, and the Resource Policy Center, Dartmouth College.

Backus, G., and J. Amlin, (1985) “Interactive Sensitivity/Confidence Analysis of Large-Scale Simulation Models,” 1985 Summer Computer Simulation Conference, Chicago, IL, July 22-26, 1985.

Backus, G. A., Schwein, M. T., Johnson, S. T., & Walker, R. J. (2001). Comparing Expectations to Actual Events: The Post Mortem of a Y2K Analysis. System Dynamics Review, 17(3), 217-235.

Barkow, J.H., Cosmides, L. & Tooby, J. (Eds.). (1992). The adapted mind: evolutionary psychology and the generation of culture. New York, NY: Oxford University Press.

Bednar, Jenna and Scott Page (2005), Can Game(S) Theory Explain Culture?: The Emergence of Cultural Behavior within Multiple Games, University of Michigan Center for the Study of Complex Systems. March 2005. www.cscs.umich.edu/~spage/culture.pdf

Boase, Jeffrey, et. al. (2006) The Strength of Internet Ties Pew Internet & American Life Project Washington,, D.C. January 25, 2006 http://www.pewinternet.org/pdfs/PIP_Internet_ties.pdf

Borum, R. (2004) Psychology of Terrorism. Tampa, FL: University of South Florida. (Monograph).

Buechler, Steven. (1995). New Social Movement Theories. The Sociological Quarterly Vol 36, Pp. 441-64.

Chambers, Lauren and Robert Glass (2005). Self-Organized Extremist Group Formation, Activation and Dissipation a Network Agent-Based Model Approach, Sandia National Laboratories.

Crenshaw, Martha (Editor) (1995), Terrorism in Context (In Introduction) Pennsylvania State University Press, University Park, PA della Porta, Donatella, and Mario Diani (1999), Social Movements: An Introduction, Oxford: Blackwell.

DeNardo, James (1985), Power in Numbers: The Political Strategy of Protest and Rebellion. Princeton, Princeton University Press.

Dennett, Daniel C. (2006), Breaking the Spell: Religion as a Natural Phenomenon, Penguin Group, London.

Doyle, J. K., Ford, D.N., Radzicki, M.J., Trees, S.W. (2001) "Mental Models of Dynamic Systems" *Encyclopedia of Life Support Systems*. EOLSS Publishers. http://www.eolss.co.uk/ and http://www.wpi.edu/Academics/Depts/SSPS/Research/Papers/27.pdf .

Drummond Jonathan T. (2002). From the northwest imperative to global jihad: social psychological aspects of the construction of the enemy, political violence, and terror, In The psychology of terrorism, by Chris E Stout Westport, CT: Praeger.

Engle, R.F., and C.W.J. Granger. (1991). Long-Run Economic Relationships: Readings in Cointegration, Oxford University Press, Oxford, UK.

Epstein, Joshua M. (2002), "Modeling civil violence: An agent-based computational approach", Proceedings of the National Academy of Sciences, Vol. 99, Suppl. 3, May 14, 2002, and is available at http://www.pnas.org/cgi/content/abstract/99/suppl_3/7243

Florens, Jean-Pierre & Hendry, David F & Richard, Jean-Francois, 1996. "Encompassing and Specificity," The Journal of Econometric Theory, Cambridge University Press, vol. 12(4), pages 620-56.

Gigerenzer, Gerd (1993), "The Superego, the Ego, and the Id in statistical reasoning", Ch. 11 in A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues, G. Keren and C. Lewis eds. Lawrence Erlbaum Associates, Mahwah, NJ

Glymour, Clark, and Gregory Cooper (eds). (1999).Computation, Causation, and Discovery, MIT Press, Cambridge, MA,

Griffiths, Thomas L. and Joshua B. Tenenbaum (2006), “Optimal predictions in everyday cognition,” Psychological Science (in press, 2006).

Gurr, T. R. (1970) Why Men Rebel, Princeton: NJ, Princeton University Press

Helson, H. (1964) Adaptation-level theory; an experimental and systematic approach to behavior. New York: Harper & Row

Hendry, David F. (1993). Econometrics: Alchemy or Science?, Blackwell Publishers, Cambridge, UK.

Hendry, David F. (1995). Dynamic Econometrics, Oxford University Press, Oxford, UK.

Hogarth, R. (1987). Judgment and choice: The Psychology of Decision (2nd edition.). New York: Wiley.

Howson, C., and P. Urbach, Scientific Reasoning: a Bayesian Approach. Open court, Chicago, 1993.

Kahneman, D. (2003a). A Perspective on Judgment and Choice: Mapping Bounded Rationality. American Psychologist, 58(9), 697-720.

Kahneman, D. (2003b). A Perspective on Judgment and Choice: Mapping Bounded Rationality. American Psychologist, 58(9), 697-720.

Keeney, R. L. and Raiffa, H. (1976).Decisions with Multiple Objectives, John Wiley & Sons, New York NY.

Lacquer, Walter (2004), The Terrorism to Come, Policy Review no 126, Hover Institution, Washington, DC.

Lazarus, R. S. and S. Folkman (1984). Stress, Appraisal and Coping. New York: Springer static

Maslow, Abraham H. (1962). Toward a Psychology of Being, New York, D. Van Nostrad Company.

McCauley, Clark (2002), Psychological issues in understanding terrorism and the response to terrorism, In The psychology of Terrorism by Chris E Stout Westport, CT: Praeger.

McCauley, Clark and Emmanuel Karagiannis (2006). Hizb ut-Tahrir al-Islami: Evaluating the threat posed by a radical Islamic group that has opposed violence for fifty years (working paper)

McFadden, D. (1982), “Qualitative Response Models,” in Advances in Econometrics, Ed. Werner Hildenbrand, Cambridge University Press, New York.

McKay, M. ,W. Conover, and D. Whiteman. (1976). "Report on the Applications of Statistical Techniques to the Analysis of Computer Codes", Los Alamos National Laboratory, LA-NUREG-6526-MS, 1976

McKim, Vaughn R., and Stephen Turned (eds). (1997) Causality in Crisis?: Statistical Methods and the Search for Casual Knowledge in the Social Sciences. University of Notre Dame Press, Notre Dame, IN.

McCormick, Gordon H. (2003), Terrorist Decision Making, In Annual Review of Political Science Vol. 6: 473-507

Meadows, D. L., W. W. Behrens III, D. H. Meadows, R. F. Naill, J. Randers, E. K. O. Zahn. (1974). *Dynamics of Growth in a Finite World*. Cambridge MA: Productivity Press.

Meadows, D.H. (1980), 'The unavoidable a priori', In: J. Randers (ed.), Elements of the system dynamics method, MIT Press: Cambridge.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Moos, R. H. (2002) The Mystery of Human Context and Coping: An Unraveling of Clues. American Journal of Community Psychology, Vol. 30, No. 1, pp. 67-88.

Naill, R. and G. Backus. (1977). Evaluating the National Energy Plan. Technology Review, July/August, 51–55.

Olson, Mancur (1984). Rise and Decline of Nations: Economic Growth, Stagflation, and Social Rigidities, New Haven: CT, Publisher: Yale University Press

Olson, Mancur (2000). Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships, New York: NY, Basic Books

Palfreyman, N. (2003) "Wicked problems and policy resistance." Weihenstephan University of Applied Sciences, Lindau, Germany, www.fh-weihenstephan.de/bt/bi/lehre/LO/pdf/LO03.pdf

Pearl, Judea. (2001). Causality: Model, Reasoning, and Inference, Cambridge University Press, Cambridge, UK.

Pinker, S. (2002) The Blank Slate: the Modern Denial of Human Nature. London, Allen Lane,

Ritchey, T. (2005); Wicked Problems. Structuring Social Messes with Morphological Analysis. From the Swedish Morphological Society http://www.swemorph.com/ and http://www.swemorph.com/pdf/wp.pdf

Rittel, H., and M. Webber (1973); "Dilemmas in a General Theory of Planning" pp 155-169, Policy Sciences, Vol. 4, Elsevier Scientific Publishing Company, Inc., Amsterdam, 1973.

Runciman, W. G. (2000) The Social Animal. Ann Arbor, MI: University of Michigan Press

Sageman, Marc (2004) Understanding Terror Networks, University of Pennsylvania Press, Philadelphia

Salmon, W. (1998). Causality and Explanation, Oxford University Press, Oxford, UK.

Simon, H.A. (1957). Administrative Behavior (2nd ed.). New York, NY: Macmillan

Schelling, Thomas C (1960), The Strategy of Conflict, Harvard University Press, Cambridge.

Silke, Andrew (Editor) (2003) Terrorists, Victims and Society: Psychological Perspectives on Terrorism and its Consequences (Wiley Series in the Psychology of Crime, Policing and Law) Wiley, John & Sons, Incorporated, Chichester, England.

Spirtes, Peter, Clark Glymour and Richard Scheines, Causation, Prediction, and Search, Springer Verlag, New York, NY, 1993

Sterman, J. D. (1988a) "Modeling the formation of Expectations: the History of Energy Demand Forecasts," International Journal of Forecasting, v.4, p.243-259.

Sterman, John D. (1988b). "A Skeptics Guide to Computer Models." pp. 133-169. In: Lindsey Grant, ed. Foresight and National Decisions: The Horseman and Bureaucrat. Lanham, MD: University Press of America. http://web.mit.edu/jsterman/www/Skeptic's_Guide.pdf

Sterman, J. D. (1989). Misperceptions of Feedback in Dynamic Decision Making. Organizational Behavior and Human Decision Processes 43 (3): 301-335.

Sterman, John (2000). Business Dynamics: Systems Thinking and Modeling for a Complex World, McGraw-Hill/Irwin, Boston

Sterman, John D. (2002). All Models are Wrong: Reflections on Becoming. a Systems Scientist. System Dynamics Review 18(4): 501-531. http://web.mit.edu/jsterman/www/All_Models_Are_Wrong_(SDR).pdf

Suppes, P. (1970) " A Probabilistic Theory of Causality:" in Acto Philisophica Fennica, XXIV, Amsterdam, North Holland.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

Weisbuch, G., G. Deffuant, et al. (2005). "Persuasion dynamics." Physica A: Statistical Mechanics and its Applications 353(1-4): 555.

Westen, Drew, (2006) "Democrats and Republicans Both Adept at Ignoring Facts, Study Finds" news release from LiveScience on paper to be presented at the Annual Conference of the Society for Personality and Social Psychology, January 26-28, 2006, in Palm Springs, California, http://www.livescience.com/othernews/060124_political_decisions.html

Wilson, Edward O. (1975), Sociobiology: The New Synthesis. Boston, Harvard University Press.

Zimbardo, P. G., Snyder, M., and Hirschman, A. (1973). Social determinants of group loyalty, innovation, and activism (ONR Technical Report: Z-11). Washington, DC: Office of Naval Research. http://www.prisonexp.org/

Appendix 1: Bayesian Causal Discovery

Causal discovery is relatively new and remains controversial.[8] The development of the technique depended in part on the greater computation capabilities available today for solving NP-hard problems via exhaustive search. The original (SGS algorithm) approach was developed by Spirtes, Glymour and Scheines in 1993.[9] It goes beyond Granger Casualty to actually claim knowledge of the causal relationship within a specified degree of confidence.[10]

In general, the approach assumes only non-experimental data -- the only data generally available in social-behavioral modeling. It combines several concepts and assumptions that may, at first, sound overly restrictive, but they are actual broadly applicable to most real-world situations.

The basic concept is to first assume actual causal relationships exist. Second, it assumes that the data available is usable to support a position on the confidence in a causal assertion. All potential causal relationships (even those that will later prove to be independent) are collected. These are assembled into all possible Directed Acyclic Graphs (DAGs). "Acyclic" indicates that no feedback loops are modeled. This does not mean that there are no feedback loops. It means that the loop contains a (mathematical) integration. The time subscript change occurring at the level allows an unfolding of the loop for statistical analysis purposes, such that the t-1 term is at one end and the "t' term is at the other of the DAG.

If we hypothesize X causes Y, and that Y causes Z, the DAG would simply be as shown in Figure 1.

$$X \rightarrow Y \rightarrow Z$$

Figure 1. Simple DAG.

Just like a flow diagram, the arrow direction indicates (assumed) causality. The causality can be a priori designated as positive or negative as a restriction to the statistical analysis of the data that determines the correlations and rejects many DAGs.

There are generally many variables and thus, many potential combinations with multiple connections, such as shown below:

[8] McKim, Vaughn R., and Stephen Turned (eds), Causality in Crisis?: Statistical Methods and the Search for Casual Knowledge in the Social Sciences. University of Notre Dame Press, Notre Dame, IN, 1997.
[9] Spirtes, Peter, Clark Glymour and Richard Scheines, Causation, Prediction, and Search, Springer Verlag, New York, NY, 1993
[10] Pearl Judea, Causality: Model, Reasoning, and Inference, Cambridge University Press, Cambridge, UK, 2001 and Salmon, W., Causality and Explanation, Oxford University Press, Oxford, UK, 1998

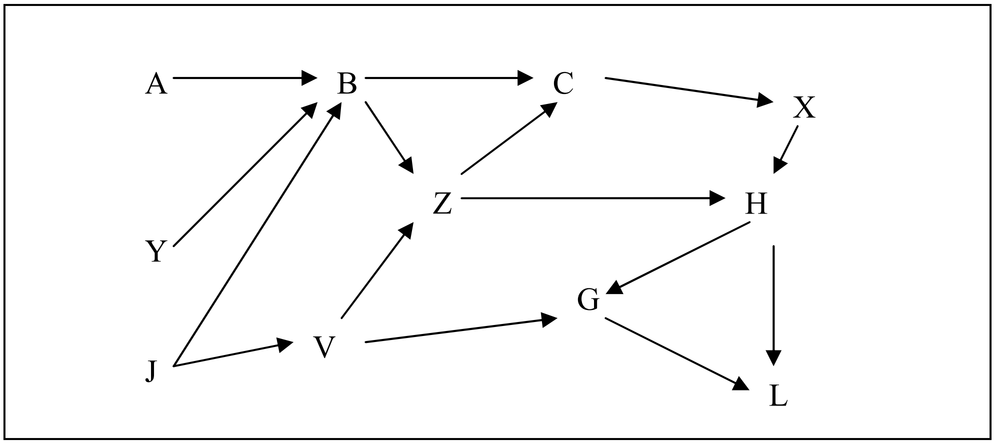


Figure 2: Multi-path DAG.

The SGS algorithm would look at all possible combinations of connections of the variables in Figure 2 as a possible cause of any other variables. The SGC algorithm looks for "d-separation" points in the DAGs it generates. In the case of "L" in Figure 2, conditional on V, Z, and X above, only G and H affect L. In Figure 1, Y d-separates X and Z, Given X, Z only depends causally on Y.

These d-separated sets are then tested to determine the correlation (or lack thereof) among the variables in a given direction of causality. If two variables are not d-separated, they are conditionally dependent. All the conditionally dependent variables (d-connected) affecting another variable in a DAG, constitute the potential casual relationship with that variable.

The data must be "faithful" to any causal assertion. The "causal faithfulness" is a key concept in that it not only implies that the representation is valid, but that there is acceptable confidence (faith) that the relationship is truly causal. All the relationships are represented as Bayesian networks where the data is used to determine the conditional probability of one variable given the d-connected variables, conditional on the d-separated variables.[11] The independence of a variable is provided by the Causal Markov condition. To have a Casual Markov condition, a variable X must be independent of (not causally related to) every other variable, conditional on all of it direct causes. That is, it must be causally faithful. Causal Markov conditions are the basis for the assumed causality found via the statistical search.[12]

Those DAG that cannot be causal due to time lag considerations (Y_{t+1} can never cause Z_t), are thrown out. Any a priori considerations that exclude some combinations are also used to reduce the possibilities that must be tested. Other prior considerations (K) are added as appropriate (or justified) and combined with the data (O) to produce the posterior probability that, for example, Y causes Z (as in Figure 1) given observations "O" and prior knowledge "K." Let "S" be the set of all possible combinations of DAG available related to Y causes Z.

[11] Howson, C., and P. Urbach, Scientific Reasoning: a Bayesian Approach. Open court, Chicago, 1993.
[12] Suppes, P. " A Probabilistic Theory of Causality:" in Acto Philisophica Fennica, XXIV, Amsterdam, North Holland. 1970.

The probability the Y causes Z conditional on O and K is then:

$$P(Y \rightarrow Z \mid D,K) = \sum_{S:\ \{Y \rightarrow Z\}\ \text{element of}\ S} P(S, O \mid K) \qquad 15)$$

Glymour provides the details of the process.[13] Statistics can further be developed to quantify the probability that the probability of causality is meaningful. These probabilities can be ranked. Usually, only one DAG stands out as most probable. Remember that by definition, all causal possibilities are present, so the result is definitionally determining "THE Causality." It is possible that there will be a small group of causal inferences with statistically comparable probabilities. Frequently, this means that all are part of a more elaborate single causal process that includes latent (unmeasured) variables not known or not available in the data. To an econometrician, this phenomenon is called encompassing. It is implied that the latent variable has interactions with all the "competing causality" DAGs. The inclusion of a meaningful latent variable can also be used to tie the set together for statistical parameterization of the equations.

Limitations of Causal Discovery

Sometimes the data are simply not available to determine the causality. This is more true for aggregate country data than company data. A latent variable that represents a critical level tightly tied in feedback violates all the premises required for causal discovery. Luckily, such a hidden level is uncommon. The technique only avoids controversy when it is used for verification of assumed causality or to sort through competing causal hypotheses during model construction. Its marginal value may be low in a consulting situation. It is still a very divisive topic among statisticians.

Value of Causal Discovery

Causal discovery techniques can add credibility to contentious causal assumptions. It can sort through contradictory casual assumptions. It can find (potential) causal relationships when the set of relationships used in the model seem inconsistent with historical data. Conversely, it can show where there is variable independence despite claims by clients of cooperative or coordinated activities. The use of Causal Discovery is then to "prove" that there is not causality. In this situation, only the DAG assumed by the client is needed to show the lack of correlation (i.e., the lack of a causal Markov condition).

[13] Glymour, Clark, and Gregory Cooper (eds), Computation, Causation, and Discovery, MIT Press, Cambridge, MA, 1999.

Appendix 2: Cointegration

Cointegration was first conceived by Clive Granger[14] in 1981, but the development of the method was not achieved until 1987.[15] He received a Nobel Prize for his pioneering work in 2003. It is now a widely accepted and used technique.[16] Cointegration focuses on determining no-memory, short-term memory, and long-term memory dynamics within a data generating process (DGP). The concept of "no memory" corresponds to algebraic equations having no time dependency. Short and long term memories are naturally associated with numerical integration levels. Short-term memory conveniently corresponds to reinforcing loops and long-term memory to balancing loops within feedback models.

Cointegration discusses dynamics in terms of variables being jointly integrated (or differenced). To economists, the effort is to find stationarity in the residual error term. That is, they want the variables to stay related without the error term growing over time. (This implies that there is no remaining information regarding behavioral.) Differencing of a series with serial correlation will always result in stationarity if differenced enough times. An un-differenced equation is designated I(0), a first difference I(1), etc.

A general introductory cointegrated equation would be:

$$\Delta Y_t = B0 + \sum Bi * \Delta X_{i,t} + Bn * (Y_{t-1} - \boldsymbol{F}(X_{i,t-1})) + u_t \qquad 1)$$

Where the Δ is the difference operator: $\Delta X = (X_t - X_{t-1})$. A problem with econometrics is that the time interval always coincides with the data collection interval. In the "DT" limit, the SD equations are differential equations. The distinction between a "true" difference equation and the 'true" differential/integral equation will result in an anomalous Δ term that is "without economic interpretation" in the cointegrated equation. The $\boldsymbol{F}(X_{i,t-1})$ is the asymptotic value of Y when X is held constant. A non-zero B0 distorts this definition and is often restricted to a 0.0 value. The $Bn * (Y_{t-1} - \boldsymbol{F}(X_{i,t-1}))$ term is called the Error-Correction Mechanism (ECM). An example of $\boldsymbol{F}(X_{i,t-1})$ could be

$$\boldsymbol{F}(X_{i,t-1}) = A0 + A1 * X1 + A2 * X2 \ldots \qquad 2)$$

The use of a non-zero B0 means that the A0 in equation 2 becomes A0'=A0+B0/Bn. The u_t is the classical "error term." Cointegration ensures that it is never serially correlated.

[14] Granger, CWJ, "Some properties of time series data and their use in econometric model specification," Journal of Econometrics, Vol. 16, 1981, pp. 121-130

[15] Engle, R.F., and C.W.J. Granger, Co-integration and error correction representation, estimation, and testing, Econometric, Vol. 55, pp 251-276, 1987.

[16] Hamilton, James D., Time Series Analysis , Princeton University Press, Princeton, NJ, 1994 and Engle, R.F., and C.W.J. Granger, Long-Run Economic Relationships: Readings in Cointegration, Oxford University Press, Oxford, UK, 1991 and Maddala, G.S. Introduction to Econometrics, Macmillan Publishing Company, New York, NY, 1992 and Hendry, David F., Econometrics: Alchemy or Science, Blackwell Publishers, Cambridge, UK, 1993 and Hendry, David F., Dynamic Econometrics, Oxford University Press, Oxford, UK, 1995.

The determination of cointegration is based on the concept of a unit root. Starting with a simple autoregressive equation:

$Y_t = \rho Y_{t-1} + \varepsilon_t$ 3)

Where ε is the error term but, unlike "u," it might be serially correlated.

If ρ is greater than unity, there is a positive feedback situation. If it is less than unity, there is a negative feedback situation.

This makes sense by rewriting the equation to look a bit more like those used in System Dynamics (SD) methods:

$Y_t = Y_{t-1} + dt*(\rho-1)*Y_{t-1}$ 4)

$\Delta Y = dY = (\rho-1)*Y = \alpha Y$ 5)

The sign of α determines the loop polarity. The polarity depends on the value of ρ compared to unity. If α is thought of as a positive growth rate, as in population growth, then ρ is greater than 1.0. The equation is not cointegrated. It only has short term memory. The level changes slowly as other possible inputs affect it. A simply SD Smooth (with its long-term, cointegrating memory) will clarify the unit-root significance.

$Y_t = Y_{t-1} + dt*(S_t - Y_{t-1})/T$ 6)

Or

$\Delta Y = (S_t - Y_{t-1})/T$ 7)

Where S in the input variable to be smoothed and T is the averaging time.

In general cointegration terms:

$\Delta Y_t = B0 + B1*\Delta S_t + B2*(S_{t-1} - Y_{t-1})$ 8)

This looks like the original cointegration Equation 1 above.

B2-1/T 9)

B1=(ρ-1) 10)

The (ρ-1) is comparable to its use in Equation 4 above.

Note the minor difference versus differential equation issue of the time subscript change on "S." This is not statistically significant, but it does change the causal interpretation of estimation results.

The regression of Equation 8 corresponds to Equation 6 or 7 only if ρ is unity – the unit root. The unit root indicates the reinforcing loop limit. The ρ just needs to be unity from a statistical perspective. A value of less than unity will do as well in most cases. There is a problem if ρ is much above unity. The test that ρ is statistically unity, is then not the conventional "t" statistic for (ρ-1) being non zero, but rather a modified distribution that is heavily skewed toward value below zero. The verification of the unit root is called the augmented-Dickey-Fuller test (ADF), in cointegration jargon. A Smooth function is perfectly cointegrated.

A population growth equation is not cointegrated. The serial correlation of the error term can be removed by simply assuming a growth rate. A growth rate equation has the Bn of Equation 1 equal to 0.0 and the Bi not all equal to 0.0.

The explicit use of lagged values determines "causality." In cointegration, the test (Granger Causality) is not to prove causality, but to verify when there isn't causality. If Y_t is a well correlated function of $X_{i,t-1}$, the X_i could be causing Y, but if Y_t is more correlated with a function of $X_{i,t+1}$ (note the "+"), then the X_t does not Granger-Cause Y. Another perspective on Granger causality is to say that Y is explained better by lags of X than by lags of Y alone.

The test of whether Y_t is a function of $X_{i,t}$, occurs in the first pass of the 2 stage cointegration regression process. The first stage estimates the long-term (asymptotic) solution and second stage estimates the dynamic ΔX contribution. Note that higher order ΔY [I(n)] components can also be added to Equation 1.

One economic breakthrough of cointegration is that price is a level.[17] SD expects that all relevant economic processes need to be represented by a feedback process. The implies the existence of a "level" and to cointegration theory that implies "memory." If the process is balancing (negative feedback), it must be cointegrated.

When cointegration methods are applied to price as a function of Supply (S -production capacity), demand (D) ,and inventory (I), the first approach is to say that:

$$P=B0+B1*S+B2*D+B3*I+\varepsilon \qquad 11)$$

If these terms are taken as logarithmic, the equation can also be written in the more common form of:

$$P=\theta*S^{\alpha}D^{\beta}I^{\gamma} \qquad 12)$$

[17] Hendry, D. "Explaining Cointegration Analysis: Part 1," The Energy Journal, International Association of Energy Economists, Cleveland, OH, Vol. 21, No 1, pp. 1-42 and Hendry, D. "Explaining Cointegration Analysis: Part 2," The Energy Journal, International Association of Energy Economists, Cleveland, OH, Vol. 22, No 1, pp. 75-120.

Equations 11 and 12 have expanding errors over time (primarily due to the inventory term). If S, P and I are turned into indices, the θ=(exp(B0)) is the "normal" price. The α and β then have meaningful behavioral or qualitative choice interpretations. The inventory term, so far, has no economic interpretation.

The cointegration version of Equation 12 becomes:

$$\Delta P = B1*\Delta I + B2*\Delta S + B3*\Delta P + B4*(A0 + A1*S_{t-1}/D_{t-1} - P_{t-1}) \quad 13)$$
$$\Delta I = S_t - D_t \quad 14)$$

Now all terms have a physical/economic meaning and generally match the data.

Most models can be represented as a Vector Autoregression (VAR) system. In a two variable system, if the characteristic root contains two unit roots, then there is no cointegration (it is *generally* a reinforcing loop). If there is one unit root, there is cointegration (it is balancing), and if there are none, it is a stationary (algebraic) equation. With higher order systems, there is a need to decompose the system in the loops and compare the unit roots for each subsystem (loop).

Limitations of Cointegration

Cointegration equations are often estimated in a logarithmic form. The dynamic behaviors then have no real-world counterpart or causal meaning and they are asymmetric when viewed in the raw units of measure. For example, second order systems do not produce sinusoid oscillations. The equations need to be linear (or log-linear) for legitimate statistical analysis. Non-linear maximum likelihood estimation allows non-linear equations, but interpretation of validity becomes more of an art form. Cointegration produces incredibly accurate forecasts in timing and magnitude. The equations are often found by using all the information available and include higher order differencing (all the combinations). This is called a-theoretical modeling. Some parameters can be construed to have an economic interpretation. Those that have none are still deemed legitimate by economists because they add to the accuracy of the forecast – and of policy impacts. The use of the discrete lagged variables means that the Δ terms are often just picking up the error in the regression of the ECM term. (i.e., the subscript problem noted earlier). The parameters may have no causal meaning to a valid causal-feedback model. The ECM term has only one time constant. A agent-based approach would assume that each mechanism affecting negative feedback would have it own, possibly dynamic time constant.

Value of Cointegration

It appears that cointegration can help with system definition and its verification. It does help determine the "goal" of balancing loops. It can determine if positive or negative feedback loops are missing. Cointegration tends to show that historical data *does not* have structural breaks. Humans make decisions as they always have, despite wars and stock-market bubbles. It indicates that the model equations should produce the "break" response from the changing input data (or endogenous model dynamics), and not from changing model structure. Dynamic models with feedback are definitionally collections of cointegration equations. The equations are not a-theoretical. They can include non-linear loop topologies that can only be approximated (linearized) in cointegration terms. Dynamic agent-based models are then definitionally more capable of providing accurate forecasts than the cointegration models. Cointegration is a well accepted part of the economic community.

Appendix 3: Expectation Formation

Expectations are fundamental to decision-making. Expectations form from a comparison of remembered history and perceived, current conditions. The perceived, current condition is called the realized condition (value) and it typically contains significant filtering to remove short-term noise or to discount anomalous events. The noise is removed via a smoothing process over a Short-term Averaging Time based on the first-order Erlang distribution common to most human activities (Griffith 2006). Similarly, the Realized Conditions accumulate to become the Remembered Conditions over a Historical Averaging Time that weights the near-term conditions more than those long past.

Humans, and all other biological system, have great difficulty dealing with absolute values. It is much more natural (and biological) to recognize relative values – such as 20% off a price or a 50% increase in income or food production. The growth or Rate of Change comes from a direct comparison of realized to historical values. Projecting this trend off the Realized values produces the Expected value either for the present or for some future period (Forecast time > 0) per the equations below.

$$RealizedValue = \int_{\tau 0}^{T} (CurrentValue - RealizedValue) / ShortTermAveragingTime * dt$$

$$RememberedValue = \int_{\tau 0}^{T} (RealizedValue - RememberedValue) / HistoricalAveragingTime * dt$$

$$RateofChange = (RealizedValue / RememberedValue - 1) / HistoricalAveragingTime$$

$$\begin{aligned} ExpectedValue = {} & RealizedValue * (1 + RateofChange * ShortTermAvergingTime) * \\ & \exp(RateofChange * ForecastTime) \end{aligned}$$

The above formulation explicitly captures the adaptive adjustment to breaks with historical trends.[18] The above formulation was initially developed and used by Backus for developing expectations about long-term project and programs, such as new technology development.[19] It was expanded and popularized by Sterman.[20]

---

[18] Carbone, R., and Makridakis, S., "Forecasting when Pattern Changes Occur beyond the Historical Data," Management Science, Vol. 32, No. 3, 1986, pp. 257-271.

[19] Backus, G. A., Schwein, M. T., Johnson, S. T., & Walker, R. J. (2001). Comparing Expectations to Actual Events: The Post Mortem of a Y2K Analysis. System Dynamics Review, 17(3), 217-235, and Naill, R. and G. Backus. 1977. Evaluating the National Energy Plan. Technology Review, July/August, 51–55, and Backus, George A. 1977. FOSSIL1: Documentation. DSD #86. Resource Policy Center. Dartmouth College, Hanover, New Hampshire 03755.

[20] Sterman, J. D., 1988 "Modeling the Formation of Expectations: the History of Energy Demand Forecasts," International Journal of Forecasting, v.4, p.243-259, and Sterman, John (2000) Business Dynamics: Systems Thinking and Modeling for a Complex World, McGraw-Hill/Irwin, Boston

Appendix 4: Qualitative Choice Theory

Qualitative Choice Theory (QCT) has a long history in psychology. It has only been fully developed for economic use through the work of Daniel McFadden (who won the Nobel Prize for the effort in 2000).[21] Independent of whether an individual is rational, irrational, profit maximizing, or satisficing, qualitative choice theory applies to the decision making process. It simply says that individuals make a choice based on their perception of utility in regard to those choices. QCT causes any and all information (preferences, tastes, price, time to delivery, "little voices," etc.) utilized by the individual to define a valid (or at least functional) representation of choice behavior. Like causal discovery, QCT starts with the data reflecting the conditional probability of a choice given possibly interacting, conflicting, and limited information.

Theoretically, any form of the probability distribution can be assumed. In practice, the Weibul distribution has the greatest numerical ease-of-use and has shown itself to be empirically the most likely shape of the actual distribution. The Weibul distribution is skewed to the left with a broad tail to the right. This implies that while individuals consider high "cost" options, they tend to focus on the lower "cost" (higher value) options. People do not have perfect information. A sampling of the population shows different perceptions of actual costs and personal preferences. The choice made is called Random Utility Maximization or RUM.[22] Figure 1 below shows the illustrative distribution of perceived price for three technologies (choices). While the QCT formulation can include any concept of culture, ideology, tastes, or preferences, etc, only the classical economic example is discussed here.

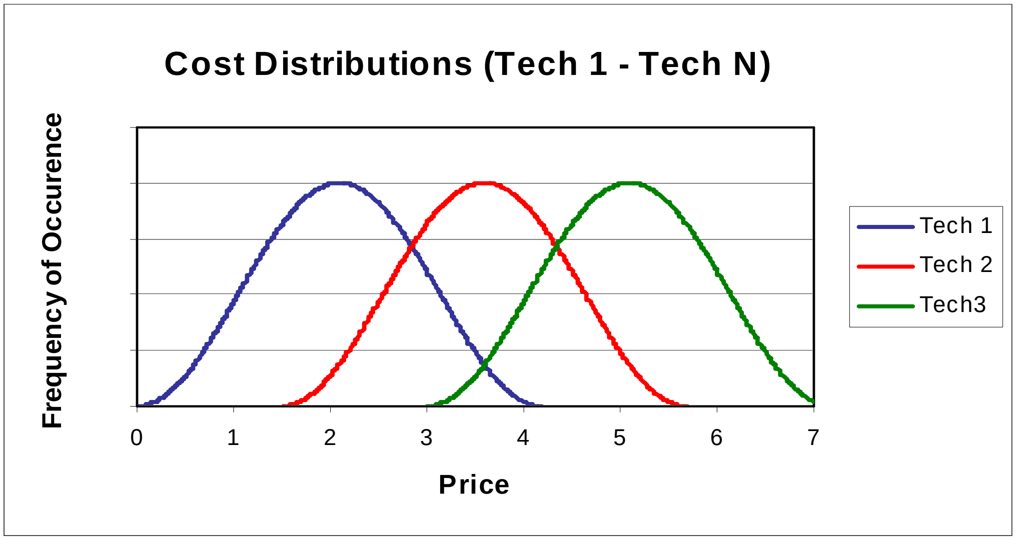


Figure 1: Illustrative Choice Distribution

[21] McFadden, D., "Qualitative Response Models," in Advances in Econometrics, Ed. Werner Hildenbrand, Cambridge University Press, New York, 1982
[22] McFadden, D., (1986), "Econometric Model of Probabilistic Choice," in Structural Analysis of Discrete data with Econometric Applications, ed. C.F. Manski and D. McFadden, Cambridge, MA, MIT Press.

Maximum-likelihood estimation (MLE) methods determine the shape of the distribution as a function of costs and preferences in the model.[23] The actual market share is determined by mathematical integration over the distributions.[24] Nonetheless, the physical process can be understood intuitively. The fraction of the time Technology 1 would be picked would be the region to the left of the red line and half the region between the left red and left green line under the blue distribution. (The half comes from the price having a 50% chance that the cost of Technology 1 is perceived as lower than Technology 2.)

Technology 2 would be selected by the fractional amount equaling one-half the area between the left red line and right of the blue line. Technology 3 would be selected by the fractional amount equaling one-half of the area between the left green line and the right blue line under the blue curve. This is the fraction of the instances that Technology 3 is perceived as having a lower cost than Technologies 1 or 2. The width (standard deviation) of the distribution can be shown to be the uncertainty in the perceived information about the technology.

The market share of Technology 1 would be as shown in Figure 2, as its price varied relative to the price of the other choices. As the price of Technology 1 becomes small compared to the other choices, its market share would go to unity. If the uncertainly is large (as in a residential decision), the slope is gradual. If there is significant effort to reduce costs (have less uncertainty), the curve is steeper as shown for industrial choices. If there is perfect information, as assumed in an unconstrained linear programming (L-P) framework, then the market share would jump from 0.0 to 1.0 with the smallest of price differentials.

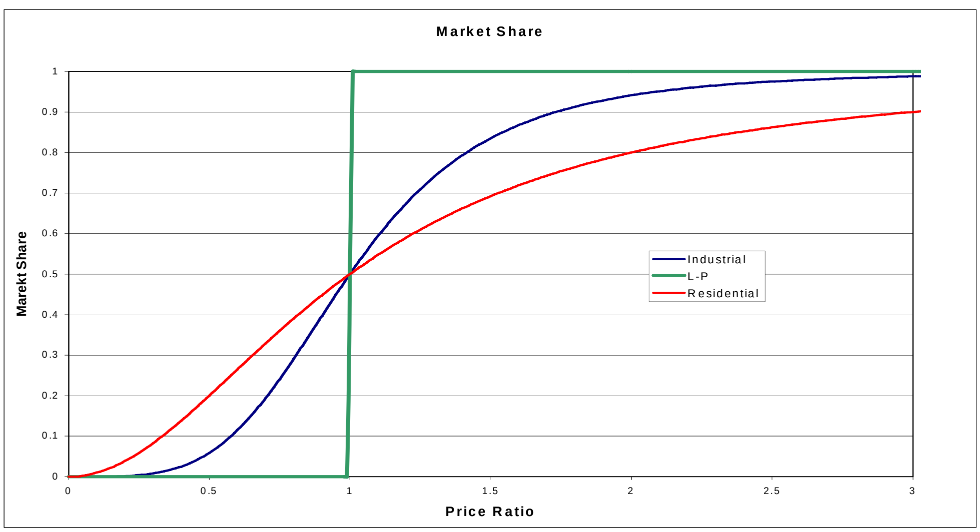


Figure 2: Illustrative Market Share Response

[23] Ben-Akiva, M., Discrete Choice Analysis: Theory and Applications, MIT Press, Cambridge, MA, 1985.
[24] McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers in Econometrics, Ed. P. Zarembka, New York, Academic Press, 1974.

The integration of Figure 1 produces the probability of the choice, or in the aggregate, the market share, of the i'th choice (MS_i) per Figure 2. For a Weibul distribution, this integral has a closed-form solution:

$$MS_i = \frac{e^{U_i}}{\sum_{j=1}^{N} e^{U_j}} \qquad 16)$$

Where U_i is the utility of choice "i," and "e" is the base of the natural logarithm.

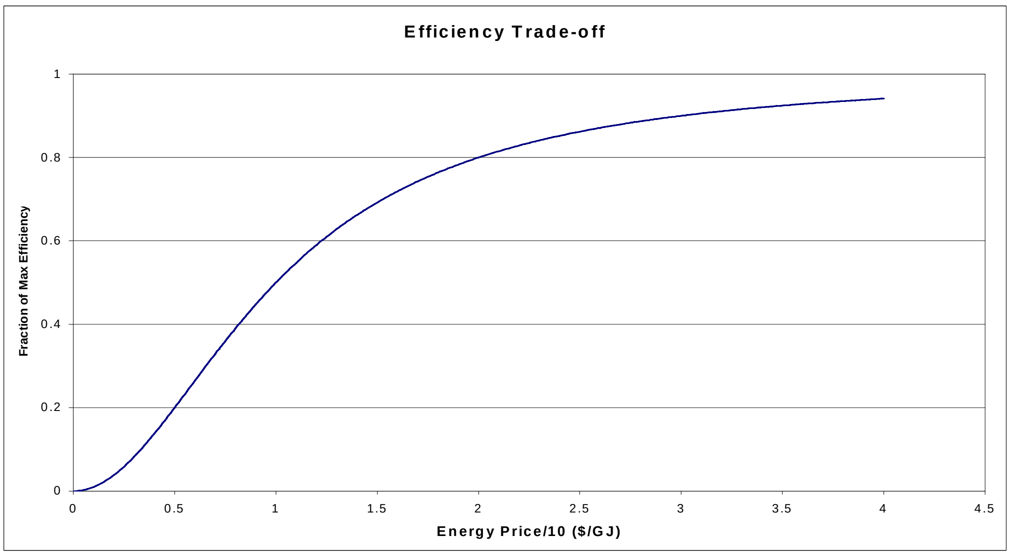


Figure 3: Aggregation of Choices.

If Technology 1 through 3 represents technology choices, then Figure 2 would represent the technology market shares on the margin. If there are many technologies, the shape of Figure 2 only changes quantitatively but not qualitatively. The sum of market shares is a market share. It is then possible to make a curve of an aggregate characteristic such as efficiency, where the choice goes from the lowest efficiency technology (when factor costs are low) to the highest efficiency (when factor costs are high). Preferences also play into this, but for simplicity, these can be thought of as added perceived costs, in this example. The curve in Figure 3 is the selected marginal efficiency at the current price and preferences for an infinite number of choices. The efficiency ratio (Efficiency/-Maximum-technological-efficiency) goes between 0.0 and the maximum (1.0). QCT allows the valid aggregation of individual choice to societal choice. This aggregation problem has haunted economics for centuries.[25]

[25] Keene, S. Debunking Economics, St. Martin's Press, New York, NY, 2001

The utility function is often written, for example, as a simple linear function of price (P_i) with the constant (non-price) term noted by Train.[26]

$$U_i = A_i + B*P_i \qquad 17)$$

In this case, the "A" would be (assumed constant) non-price factors of taste and preference for the i'th choice. It can also capture the ability to make the choice (e.g. the limitation of physician selection in an insurance plan) or the availability of the choice (e.g. the winter availability of solar energy at the South Pole). Note again that the B does not have a subscript. The "B" is directly related to the uncertainty of the choice – how well the information of the choice *set* is known and understood. The uncertainty of the decision process is the same for all choices in a set because it is an ordinal and not a cardinal process that compares all options at once.

There can be a hierarchy of choice, like a binary tree, but called nesting. Each level is a choice among all the options of that level (e.g. choosing the flavor of ice cream to eat occurs, after choosing which place to go for the snack, after the decision to go for a snack.) Each decision level is self-contained but can be conditional on the level below it.

The derivation of the theory of QCT requires that all choices at any level are mutually exclusive (e.g., the decision to live in Boston or Austin). Empirically this limitation is non-binding. A classic example is the addition of travel choice by painting half of all the buses green and the remaining buses blue. There really has been no change in the choices -- taking the green bus is no different than taking the blue bus. The "A" of equation 17 can capture this fallacy by simply multiplying the blue-bus and green-bus choice, in this example, by 0.5. The same process can often allow the complicated nested equations to be reduced to a single layer called a "comb" that requires only the single use (and estimation) of Equation 16.

Reducing the uncertainty, increasing the understanding of the choices, and making better decisions (as contained in the "B" term), takes time and effort. The benefit may not be worth the effort. When buying a house, a purchaser may want to know the price within 1% or less. For a candy bar, a 200% uncertainty variance is tolerable.. The consequences of purchasing a house are much more momentous than purchasing a candy bar. The magnitude of the "B" appears to vary directly with the importance of the decision. That importance is the cost of the decision compared to the value of the entire output (a labor-year of income for a person and the revenue for a company).

If many choices are aggregated to produce the equivalent of Figure 5, the magnitude of "B" will be reduced. This reduction occurs because aggregation is like taking a weighted average. It "smoothes" the response intensity compared to a disaggregated either-or-situation.

Data indicates the linear function of Equation 17 works well for small variations of the input variables, but the actual underlying function is logarithmic. Equation 18 is a simple logarithmic enhancement of Equation 17.

---

[26]Train, K., *Qualitative Choice Analysis,* MIT Press, Cambridge, MA, 1986.

$U_i = A_i + B*\ln(P_i)$ 18)

This indicates that people can determine relative proportionality but not absolute differences in price (or other components of utility). This implication is consistent with the previous discussion that the B is proportional to the percentage impact it has on total outcome.

If Equation 18 is substituted into Equation 16 and "m" is defined as

$m_i = \exp(A_i)$ 19)

then Equation 16 becomes

$$MS(i) = \frac{m_i P_i^B}{\sum_{j=1}^{N} m_j P_j^B} \qquad 20)$$

This equation is consistent with the engineering assessment of options according to the distribution of (estimated) cost versus (estimated) performance. The uncertainty of the estimate (the "B") is also a function of the importance of accuracy. This is the only example the author knows, where engineering theory and economic theory agree.

While MLE is required for the unbiased estimation of Equation 20, within a feedback system, ordinary least-square estimation often produces adequate parameterization to generate accurate forecasts.

Note that because the decision process is always ordinal, there is no absolute concept of preference. Therefore, one of the "m_i" must be arbitrarily selected as the numaire and set to unity.

When used over a 50 year period to simulate, for example, energy demand, some limitations of Equation 18 start to appear. One obvious area is the impact of income on decisions. A cup of Starbuck's coffee is more expensive than one from home. The ability to afford luxury items affects demand. Changes in the disposable income (I), relative to the minimum (I_m) needed to maintain health, affects buying and other decision responses.

All goods provide a service. That service is "demanded" relative to the production of output, be that output a labor year as measured in annual income units or the revenue from an industrial widget. The market share is for a service. The price is the cost per unit. That unit is a factor input to production. (Food is required to produce a labor-year; iron is required to produce a power plant.) The units of price are \$/factor-unit. For "B" to have a probabilistic meaning, the "P" term must be a proportion. The proportion is the comparison of the factor price to the price of output. That process requires a conversion term whose units are \$ of output per factor unit, here defined as the economic intensity

EI. EI is the measure of efficiency in using the input factor. With the improved income and price concepts, equation 18 becomes:

$$U_i = A_i + B*\ln(P_i/EI)+C_i*(I/I_m) \qquad 21)$$

The term P/EI has the units of $-factor/Factor-Unit / $-Output/Factor-Unit. The factor-units cancel, but the dollars do not. The dollar units of measure have important adjectives. The rigorous use of QCT could lead to the claim that there is no such thing as a dimensionless number and that valuable interpretive/causal information is lost by attempting to use dimensionless numbers. (As another example, energy efficiency is not a canceling Btu/Btu ratio, but rather BTU-Service-out/Btu-Primary-in.)

The Income term of Equation 21 is adequate for simulating demand in industrialized countries, where I/I_m is significantly greater than unity. A more complicated formulation is needed for values below and near unity.

The "A" term can be divided into its separate components ($A_{i,k}$). One term will always have to contain "other" residual "A" components. For policy purposes, these components can represent advertising, availability, color, style characteristics, or anything else that might affect the choice. Care is needed during estimation to avoid spurious parameters due to too many degrees of freedom given the quality of the data. Equation 21 then becomes:

$$U_i = \sum_k A_{i,k} + B*\ln(P_i/EI)+C_i*(I/I_m) \qquad 23)$$

Note that because the "$A_{i,k}$" are use in the exponential context of Equation 16, they reflect relative, rather than absolute preferences, From a QCT sense, they are proportional just like all other terms.

Limitations of Qualitative Choice Theory

Non-price terms may be truly co-dependent. Peer pressure and the reduction of early-adopter risk may be both a function of knowledge of other users (i.e., the average existing market share). It can become difficult to break out the separate influences.[27] Naïvely defining decision components can lead to invalid conclusions.

Value of Qualitative Choice Theory to SD

The use of QCT seems to force a rigor and a method for defining the implicit or explicit decisions associated with a rate formulation. Experience indicates that QCT forces a self-consistency of thought and theory that always has a causal description consistent with empirical data.

[27] Keeney, R. L. and Raiffa, H., Decisions with Multiple Objectives, John Wiley & Sons, New York NY, 1976.

Appendix 5: Coping-Capacity as an Explanation Of Individual and Societal Performance

George Backus
President, Policy Assessment Corporation
Denver, Colorado

October 22, 1999

Abstract

This paper describes simulation research on the dynamics of coping-skills. Based on the work of previous researchers, this paper argues that the concept of coping-skills is applicable to analyzing and understanding the dynamics that allow individuals, firms, and societies to survive and prosper as the environment around them or within them continuously changes. While the equations are included here to support the theory, the focus of the paper is on the apparent broad applicability of the approach using examples. The coping skills to respond to stresses changes dynamically in relation to the past stresses and associated past responses. Several examples of the evolving coping skills and response dynamics are presented.

1. Introduction

Coping-skills is a concept that cannot be narrowly defined. In broad terms, a coping skill is an ability of an individual, a family, a group, or a society to respond successfully to environmental stimulus. Here stimulus is defined in terms of perceived psychological pressure. These pressures can be internal as well as external. The internal pressures, however, are the probable consequences of an interactive accumulation of responses by an individual to external pressures. Thus, a coping-skill is, for the most part, a *learned* characteristic. The concept of coping skills applies equally well to plants and animals, but since humans appear to have a much larger range of coping-skills, the focus here is on human coping skills, including the composite coping-skills of human organizations such as firms.

The concept of coping-skills is illusive because a one-to-one relationship exists only between the most narrowly defined stimulus and response. For example, a firefighter may routinely respond to life-threatening events, but be unable to give a speech in front of a group. A stand-up comic who lacks musical aptitude and who is put on the spot to play the piano at a formal party is well beyond his/her coping-skills relative to music, but still may be able to use interpersonal skills to make the event enjoyable for all. Does the stand-up comic necessarily have better coping skills than the firefighter? All that can be said is that each individual has multiple attributes, each embodying a coping-skill. The complexity of most interactions often allows an individual to substitute a stronger coping-skill for a weaker one to maximize the successful *counter*-response (often as a counter-force) to a perceived (often as a threatening) situation.

From this perspective, coping-skills are a critical determinant of psychological well-being. They are strong indicators of how an individual, family, society, or firm will function and, hopefully, prosper in response to uncertain and changing events. Coping skills determine an individual's ability to learn, to be motivated, to cooperate, to react, to create, to cope, and to survive. In the extreme, it is possible to argue that coping-skills are the foundation of all human behavior (Moos, 2002). All aspects of life result in and from making choices. Those choices are solely the consequence of an individual's ability to perceive a condition and to choose as the individual's situation allows. That "situation" and the associated weighing of any alternatives embody the essence of coping-skill dynamics.

A coping-skill is an operational characteristic loosely restricted by evolutionary limitations. Individuals have certain naturally endowed capabilities and limitations that reinforce or discourage specific activities. For example, an individual with an inner-ear (orientation) dysfunction would tend to avoid a career as a test pilot. An individual with a significant hearing loss may avoid conversation and lose the ability to interact successfully with others. Just as an individual is limited, a department in a firm, or the firm itself is limited by its in-place process-equipment and financial strength, a society is limited by its in-place infrastructure and its financial (and military) strength. If these conditions are outside the control of the observed individuals, then the conditions act as immovable limitations and coping-skills must be used to extend, or maximally mitigate, those limitations. If these conditions can be changed, then coping-skills determine both the ability and activity to change them as conditions dictate.

At the individual level, innate capabilities (propensities) set the stage and coping-skills dictate how the individual acts within any setting. In a manufacturing firm, the equipment or infrastructure determines the basic functionality, and the coping-skills of the workers determine how well (and to what extent) that equipment/infrastructure will be utilized (or changed) for the benefit of the firm. The physical equipment or organizational structure dictates the capacity constraint. The coping-skills dictate the operational constraint. A change to capacity or its characteristics, in response to environmental changes, is relatively slow. This type of change takes time. Coping-skills provide the immediate response to changing conditions--although coping-skills themselves can only gradually change over time. Previous research work has looked at coping as a static process (Lazarus& Folkman, 1984). This work focuses on the dynamic aspects as being paramount to understanding behavior.

2.0 Simulating Coping Skills

Computer simulation, particularly system dynamics, can be used to test various theories and help determine those most consistent with empirical data and most useful for developing approaches that promote productive behaviors. This paper synthesizes and extracts the key concepts from simulation models of coping-skills by other researchers to develop a generic model that illustrates the implications and malleability of coping-skills (Donnadieu & Karsky, 1990; Golüke, 1980).

Understanding motivation, learning behavior, stress control, performance levels, and employee burnout, represent major challenges to employers, social worker, educators, and psychologists. Previous simulation research noted the concept of "coping-skills" as a key indicator of how people will respond in (often stress-inducing) situations (Golüke, 1983, 1984). Research based on Golüke generalized the coping-skill concept to generically include the behavior of all organisms from bacteria to humans (Backus, 1980b). The distinction between coping-skills and evolutionary responses is the difference between operational changes (organisms can accommodate different environments) and the need to change structure for survival (for example, the development of lungs for air-breathing fish).

Daily life can require an individual to use coping skills in many areas, from basic interpersonal communications to national, crisis-level decision making. The specific type of coping-skill an individual uses depends on the conditions. One type of arbitrarily defined coping skill would be used by an individual facing a stressful, but quasi-steady state condition. A quasi-steady-state condition can have either widely varying conditions within a given category like the "routine" of a swat-team member or the caseload of a social worker, or relatively constant conditions such as the activities of the stereotypical assembly-line worker. Another type of coping skill would be used by an individual facing a continuously escalating stressful condition. The escalating condition might have widely varying extremes but requires or allows individuals to continuously improve or change their skill set, such as R&D scientists and high-technology management.

For the quasi steady state condition, it is important to consider the level of effort or energy needed to perform a function. For example, how many hours can an individual stay on task? How effective can an individual be at changing the environment? Exhaustion greatly reduces operative coping-skills. An individual who lacks sleep and continually fights unchangeable conditions will wear down both physically and psychologically. While the consideration of depleting energy-levels can be added to the work reported here, those consideration have been presented elsewhere (Levine, 1985; Homer, 1985)

This paper focuses on those conditions that continuously escalate and place increasing demands on individuals. For example, how does an individual adjust to state-of-the-art technologies and the subsequent market impacts?

2.1 Attention and Response

The performance level of an individual, in response to an external or internal event, can be separated into two components. The first component is the attention that the individual assigns to the event. For example, the majority of people will walk up a few stair-steps without even recognizing the event. The sensory system of most people goes to maximum attention, however, at the sound of neighborhood gunfire. A comparison of a current situation with routine situations determines what level of attention a situation requires. This comparison can be defined as the proportional relationship between the current pressure (that is, situation) and the current level of coping-skills, as depicted in Figure 1.

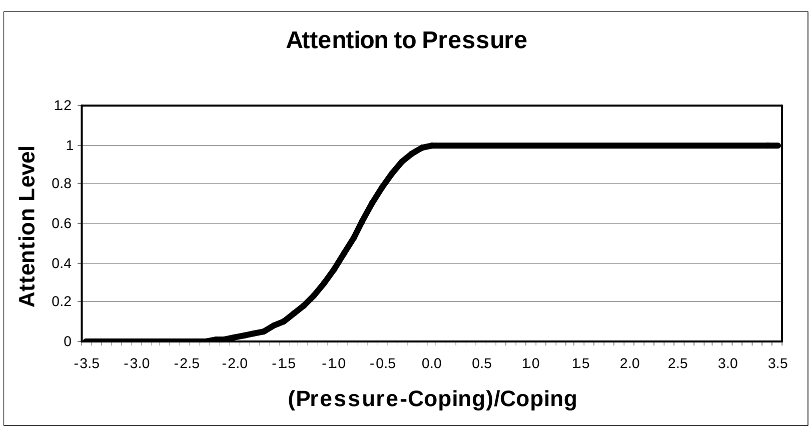


Figure 1: Attention Behavior

There are several key points to note about the construction of the curve in Figure 1. The purpose of this exercise is to build a computer simulation model to test hypotheses about coping-skills and their consequence. Like a science experiment, a simulation model cannot prove what is “truth,” but it can falsify alternative approaches and therefore give credence, or a degree of confidence, to the remaining hypotheses (Popper, 1982). The logic of the working hypothesis is then the best choice available to formulate policy. This kind of causal simulation gives social scientists a valuable tool that may help them reach a valid understanding of events and behaviors.

The mathematical process of normalization allows the quantification of arbitrarily defined concepts such as attention. Universally, any quantity can be defined in terms that bound its minimum and maximum value. A value of zero (0.0) denotes minimum value and a value of unity (1.0) defines the maximum value, as depicted in Figure 1. By definition then, the curve rises from zero to unity. The exact shape of the rising curve can best be understood by considering Figure 2.

Figure 2 shows the reaction to pressure or a stimulus that would prompt a response. For any pressure comparable to the level of coping-skill, a 100% successful response is usually achieved. If a healthy adult falls up a stairs, there is usually some extenuating circumstance that goes beyond the “routine” act of climbing stairs. In this paper, “routine” implies familiarity and a lack of associated anxiety rather than just a repetitive act.

If the pressure exceeds the “routine” ability to cope, the probability of success falls. An individual who perceives the probability of failure may respond by not trying, and by not wasting energy reserves on responses that, on average, have little or no benefit.

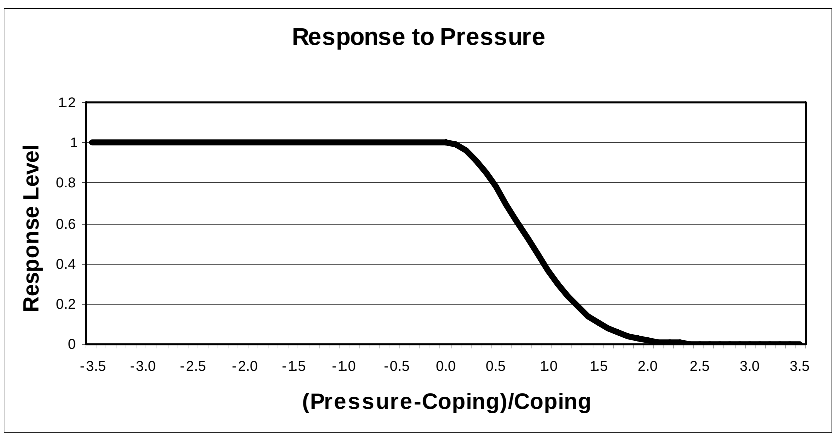


Figure 2: Response Behavior

The shape of the attention and response-curves correspond to a probabilistic concept. Individuals make choices based on their understanding of the outcomes of those choices. Both the understanding of the choices and value of the outcomes can vary around a norm and over time. People may be satisficers or may attempt to maximize the utility of an outcome (Simon, 1957; Ben-Akiva & Lerman). They may be conscious or unconscious of the choice process, or the choice may appear logical or perverse to the outside observer. This view of human behavior is consistent with and readily applied to the economic decisions individuals make, but its basis is grounded firmly in psychology (McFadden, 1986; Bock & Jones, 1968).

Attention is comparable to the conventional concept of motivation. The use of probability relative to motivation has its precedence in the work of Atkinson (Atkinson, 1964, 1966). Describing the probabilistic nature of what an individual will choose and can choose has a long history (Luce & Suppes, 1965). Luce derived a probabilistic choice approach based on the Gumbel/Weibull distributions. This approach uses what is called the multinomial logit and it accurately describes many choice situations. (Luce, 1959)

The mathematical specification of the multinomial logit is shown below where U_i is the subjective, perceived utility of choice "i."

$$M_i = e^{U_i} / \sum_j e^{U_j} \qquad \text{Eq. (1)}$$

In equation 1, the probability of selecting choice "i" (M_i) would vary between zero and unity. If there are only two choices (e.g., to do or not to do), Equation 1 takes the form:

$$M_i = 1/(1 + e^{-U_i}) \qquad \text{Eq. (2)}$$

and is called the binomial logit. To obtain Equation 2, the utility of the paired status-quo choice is defined as unity.

When viewed from the perspective of a single-event choice the equations simply represent a probability. When viewed as the representative response over many choice events, the equations represent the average extent of a response defined over a minimum (0.0) to maximum (1.0) range. the multinomial logit has been successfully applied over the full range of individual to societal choices, and it appears to be a robust approximation of most choice processes (McFadden 1986). In this work, the utility is a function of a gap (G) between the stress (pressure) and the coping skill of an individual. The larger the gap, the more negative the utility – similar how price would be considered in an economic choice. A simple expression for a utility that only includes the gap is:

$$U_i = SV * \ln(G) \qquad \text{Eq. (3)}$$

The logarithmic term indicates that stress is perceived in relative rather than absolute terms, as must be the case for information that only had ordinal meaning. SV is the sensitivity to variation and it reflects the sensitivity to the input (G in this instance) and the variance within which the input needs to be known to cause the response. The SV is negative in this case to denotes that the larger the gap, the less the utility. A SV of –2.0 (as used for this work), very roughly speaking, implies that the utility is uncertain within 50% (1/SV) of its mean value. The M function produces a sigmoidal curve going between 0 and 1 as noted in Figures 1 and 2. To allow a smooth visual representation and to avoid the need to explain inconsequential secondary dynamics in the examples presented later, the M functionally used here is:

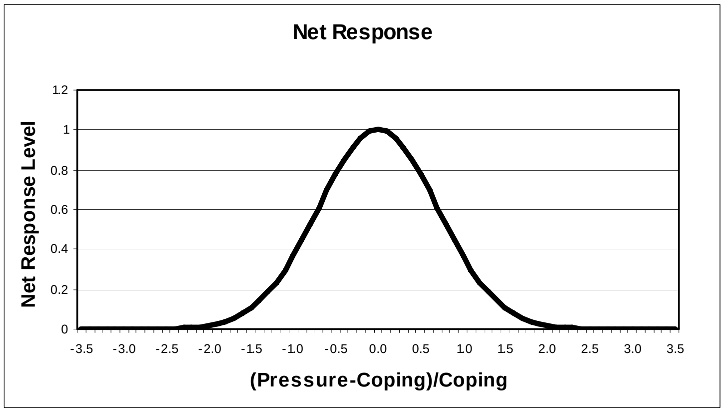
$$M_i = e^{-SV_i * G_i^2} \qquad \text{Eq. (4)}$$

This form still produces the relevant sigmoidal shape but has sharper inflection points that make the inflection points more distinct for explanatory purposes.

For the attention-behavior (a) in Figure 1, M drops from unity as the gap for the decreases below zero, slowly at first and then rapidly as the magnitude of G becomes large. For the response-behavior (r) in Figure 2, M drops from unity as the gap for the response increases above zero, again, slowly at first and then rapidly as the magnitude of G becomes large.

An individual's net response to a pressure or stimulus is the combined effect of attention and response. Being in the path of a house-sized meteor may catch one's attention but evoke little substantive response. Similarly, the stair-step example, noted earlier, produces a response but little attention. The net action (response) is defined as that action produced by simultaneously non-zero response and attention behaviors. This condition is the operating regime of an individual where coping dynamics take place. The net action behavior of attention and response is shown in Figure 3.

Figure 3: Net Active Behavior



The curves in Figures 1 and 2 capture a single mechanism. The curves are monotonic; otherwise, they would be capturing more than one mechanism. The combined result is Figure 3. The interpretation of Figure 3 would be ambiguous if it were not a composite of more basic mechanisms. The curve also uses affine definitions. There is no absolute positioning. The relationships are relative and generic. The concepts that the curves describe have no universal measure (cardinal) or meaning and, consequently, only relative (ordinal) statements comparing like concepts have any relevance. A mathematical representation that did not limit itself to monotonic and affine representations would hide implicit, and therefore, misunderstood assumptions. Early work by Hunter (Hunter et. al., 1984) attempted to apply a functional view of attitude (attention) but unit of measure and causality issues (Levine 2003) limit the applicability of that work to the approach take here.

The balance-point is that position on the net response curve that balances the pressure and the coping-skill to deal with it. This concept of balancing forces has had precedence in the field of physics since the time of Newton. The concept, however, is also relevant in the social sciences, particularly in history and psychology. Success and failure change an individual’s coping-skills. The balance-point conveniently defines the position on the net response curve where coping-skills remain stable – they produce an action adequate to counter the pressure. The actions of an individual proceed until balanced by a counter action. This is very evident in the actions of a megalomaniacal dictator who oppresses others until stopped by country borders, physical force, or finite resources. In most interactions, the actions of one individual eventually interfere with the actions of another. A conflict of some sort ensues until the pressures balance, that is, until there is no benefit to either individual in changing his/her response.

Parsimony (Occam’s Razor) is often used as a criterion to compare competing theories. Causal simulations, however, naturally show that the component mechanisms need to be simple if they are to produce the complex behaviors of real systems. Parsimony of mechanisms is a law of systems not just a philosophical tenet. If the component parts were complex, then there would be no flexibility; they could only work in one manner or produce a few behaviors. Simple parts that fit and interact with many parts provide the generic means to produce stable complexity (Wolfram, 2002). It is the composite interaction of many simple components that produces a system capable of sophisticatedly adjusting to an unlimited variety of conditions. This approach does not fall under the label of reductionism because the system produced by the parts naturally has dynamics that exceed what could be produced by any of the simple parts individually.

As individuals adjust to new experiences, they increase their coping skills. Consider this example: A high school student may find it highly stressful to write a term paper. Ten years later, on the job, however, that same person routinely prepares high-level reports and presentations. Alternatively, consider how most people have adjusted from the mechanical typewriter to the hyper-linked word processor world. Coping skills increase with successful experiences that build on earlier experiences.

Change can only occur if there is a gap (G) between the current environmental condition (pressure P) and the condition (coping-skill level CS) that minimizes the stress on the individual or entity. In the language of physics, there can only be change if the balance of forces readjusts, and the adjustment will attempt to minimize energy expenditure. Without a redefining force, inertia makes a physical body inherently maintain its current trajectory (momentum). Humans exhibit the equivalent of inertia (Backus, 1980a, pp. 28-36). They dislike change, however they may define it. Change, as used in this paper, means altering the routine conditions. For an explorer, the stressful change may be an extended waiting period before the next dangerous adventure. For a couch potato, it may be the cancellation of a weekly TV show.

The equation for the gap is shown as Equation 5 below. The CM is the maximum coping-skill. It is the biological or evolutionary limitation of an individual to cope with the environment. In absolute terms, CM is harder to specify than the concept of coping skills, but it does have intuitive meaning. For example, innate characteristics may limit the extent to which an individual can be a successful marketing representative, professional ballplayer, or a lead researcher with the human-genome project. For all but one example, CM is set to the equivalent of infinity for the simulations to be presented below. In the one example it is used, it shows the consequences of human limitations and furnishes a stronger sense of reality to the conclusions drawn in this study.

$$G = (P - \min(CM, CS * (1 + CO))) / CS \qquad \text{Eq. (5)}$$

Equation 5 looks like the title of the x-axis in Figures 1 through 3, but has some added details. The term CO is the Coping-Skill Offset, the amount of pressure it takes to cause improvement. This is the amount of external pressure beyond the routine condition needed to induce a new reaction. The term CO captures the concept that after an individual acclimates to the current conditions, the pressure no longer induces the level of response that it initially did. It simultaneously has the effect of designating how different a new pressure has to be before it evokes a change in (or threat to) existing coping-skills. CO is set to 20% for all simulations, meaning that a 20% increase over ambient pressure is needed to peak the interest of the individual and cause the maximum response shown in Figure 3.

The mathematical formulations describing the Attention and Response curves shown in Figures 1 and 2, respectively, are shown below.

$$A = e^{(-SV*MIN(G,0)^{VF})} \qquad \text{Eq. (6)}$$

$$R = e^{(-SV*MAX(G,0)^{VF})} \qquad \text{Eq. (7)}$$

These are the specific application of Equation 4. If, as a test, the magnitude of VF is made smaller, R and A would respond to small variations in G to a more limited extent. If the magnitude of VF was larger, R and A would be stable to small variations, but change quickly once G varies significantly. SV, as a measure of the variation, determines how broad the distribution shown in Figure 3 is. In concrete terms, SV defines the propensity for perfection/specialization within an individual or group of cohesive organized individuals. The greater the specialization focus, the lower is SV. With specialization, the same repetitive event continues to lead to improvement. A low SV may correspond to the perfection characteristics often attributed to performers, sprinters, or tennis pros. Similarly, this value of SV allows the individual to accommodate challenging conditions well beyond normal operating conditions. The lower the SV, the more energy is expended in return for minimal performance improvement under normal conditions and improved performance under low-probability extreme conditions. A large SV results in a successful response over a relatively wide range of stimuli, but essentially no response outside that range. This type of response is typical of the personality often termed a generalist and implies that to maximize “routine work” coping skills, individuals should experience continuously varied tasks -- except for those few individuals who need to excel in only one area.

One interesting SV dynamic is not included in this model. Totally new experiences may not have a coping-skill associated with them (other than a genetically programmed instinctive response). The individual then has no basis to be threatened by the experience and the SV is very large. Thus, children and adventurous adults are very receptive to learning from new experiences. As the new coping skill develops and interacts with existing coping skills, the SV compresses to stabilize at its long-term value. This SV dynamic contributes to societal conformance, cohesion, and solidarity, with both its positive and negative ramifications. There are limits to the productivity gains from specialization, or in the context of economists, the division of labor can only improve productivity so much.

2.2 Coping Skill Enhancement and Atrophication

The combined interaction of attention and response enhances coping-skills (CE) as stated in Equation 8. The enhancement is proportional to the existing coping-skills, attention, and response.

$$CE = (CS * (A * R))/SI \qquad \text{Eq. (8)}$$

The term SI is the Stimulus Impact and captures the maximum growth in coping-skills from an event. The value SI implies the number of times an individual must encounter and counter the pressure before coping-skills fully adjust to that pressure. The value used for SI is 5.0, which reflects a 20% adjustment per event-time. Note that the adjustment may not necessarily lead to a capability to confront the pressure; it may lead to a well-tuned avoidance behavior when the pressure far exceeds coping-skills and the response drops to nil.

Although the simulation moves through time, it is incorrect to think of the pressure as a continuous condition. Further, most “pressures” coincide with an event. The definition of pressure, as used in this paper, is the average pressure or the effective pressure caused by a sequence of events. Thus, the concept of event-time rather than just time is use to designate the time sequence.

The relative impacts of SV and SI mean they are directly related. A specialist would have a large SI and SV to capture more learning from extreme events, while a generalist would have a small SI and SV to maximize the response to pressures within the realm of existing coping-skills. For extreme events, the generalist is at a disadvantage; for routine events, the specialist wastes effort.

Coping-skills are a level of capability, a stock of resource. If an individual ceases to fully use his/her coping-skills, they atrophy. A pianist must practice throughout life, or he/she gets “rusty.” Pilots, firefighters, and nuclear-plant operators must undergo continuous re-training to maintain the required “edge” for their job. This paper considers atrophication as the primary way to lose coping skills. Atrophication (CA) reduces the coping-skills per unit of time at a rate corresponding to the Atrophication Time (AT), here set to 10 time-event-units.

$$CA = CS/AT \qquad \text{Eq. (9)}$$

Through the atrophication process, a reduced pressure level merely acts as a lower limit to which the coping-skill can atrophy. The loss of coping-skills through atrophication is equivalent to a loss of information and therefore its representation is the information delay or smoothing functions used for such purposes in the field of System Dynamics (Forrester, 1961).

Failure would seem to be a potential cause of reduced coping-skills. The probability of failure increases as the pressure exceeds the coping-skills and the successful response (Figure 2) approaches a 0.0 value. My informal review of many years’ worth of evidence within the community of technical analysts and consultants indicates no direct relationship between failure and coping-skill levels, as long as the coping skills are perceived as compatible with the pressures experienced. If the failure is due to the task being well outside the recognized coping-skill level, then personnel take the failure as simply losing a game against a bigger opponent. While also learning to avoid such situations in the future, personnel display an increase in coping skills because of the experience and "surviving" the event. Conversely, if the failure results under conditions where capability should have clearly won out, personnel soon replace the disappointment with the recognition that they can do better next time.

Punishment is another way to reduce coping-skills. This paper only briefly mentions punishment because the focus of this paper is on improving coping-skills, not destroying them. Coping skills do decrease when punishment is associated with a response (successful or failed). The effects of punishment can precipitate internal responses in an individual that perpetuate the impacts of punishment whenever the individual experiences similar conditions. Thus, it is not failure, per se, but a secondary mechanism associated with punishment that results in a loss of coping-skills. When it is critical that an individual never attempts a particular action again, such as those acts that societies often call crimes, then punishment might theoretically produce the desired results. The interaction of coping-skills, however, may cause the punishment to diminish activity in one area while simultaneously reinforcing capability in another, possibly even less desirable, area. Observers often credit the reward and hierarchical system within a firm with producing strong negative reinforcements to new ideas (turf wars) and focusing the responsibility for any "negative impacts" on a specific individual. These negative-reinforcement activities reduce the coping capability of the entire firm. This logic would appear to apply to societies as well. From the perspective of this work, "punishment responses" (or any violent reactions), to a large degree, are simply an inferior coping-skill response on the part of an individual or group of individuals who feel threatened by the actions (pressures) created by others.

2.3 The Level of Coping

The absolute level of coping-skills does not directly relate to the actual response taken. In fact, the amount of response may be inversely related to the level of coping skills, as in the case of a bully or rogue national leader. The show of force is the limited coping-skill response. Other countries must respond to counter and suppress the coping skills of the aggressor, otherwise animosities continue until a much larger force has no choice but to respond -- as was the apparent case when Hitler confronted England's Prime Minister, Neville Chamberlain in World War II. Whether considering Slobodan Miloševic, Saddam Hussein, or Josef Stalin, the logic is to early-on prevent coping skills from increasing. The process of keeping the coping-skills among competing interests balanced results in the greatest stability. This balancing limited the extent of the Cold War and many of the Middle East conflicts. Any system of interacting pressures seeks a balance. The balance may appear to be neither rational nor tolerable to any party. Scientific logic may be the strength of one group. Political acumen may be the strength of its balancing counterpart.

Because coping-skills (CS) are, in essence, a level of capability, the equation below represents them tautologically as an integration process. The coping skill enhancement (CE) compares reality to the level of coping skills to produce an adaptive response consistent with earlier research on adaptation (Helson, 1964).

$$CS = CS_0 + \int_{t_0}^{t_f} (CE - CA) * dt \qquad \text{Eq. (10)}$$

where CS_0 is the initial level of coping-skills, t_0 is the initial time the simulation starts and t_f is the time at which the simulation ends.

Although for introductory purposes the equations above have used the current pressure (P) to determine the response of the coping-skills, individuals actually respond to the assimilated pressures (AP) or the level of pressure they perceive. As in the case of atrophication, assimilated pressure is an information-smoothing process using a time constant (PAT) of 2.0 time-event units, for this work. In all the previous equations, the simulation actually uses AP instead of the previously noted "P."

$$AP = AP_0 + \int_{t_0}^{t_f} (P - AP) / PAT * dt \qquad \text{Eq. (11)}$$

Figure 4 shows the influence diagram of the coping-skill model that combines all preceding equations. The arrows show the direction of causality among the variables. The associated plus and minus signs show the reinforcing or balancing relationships of the variables, respectively. The coping skill is “boxed” to show it is an accumulated quantity while the pressure is show as being exogenous. More sophisticated versions (not reported here) can have multiple interacting coping skills responding to multiple interacting pressures. The pressures themselves can be the coping skills of multiple interacting entities.

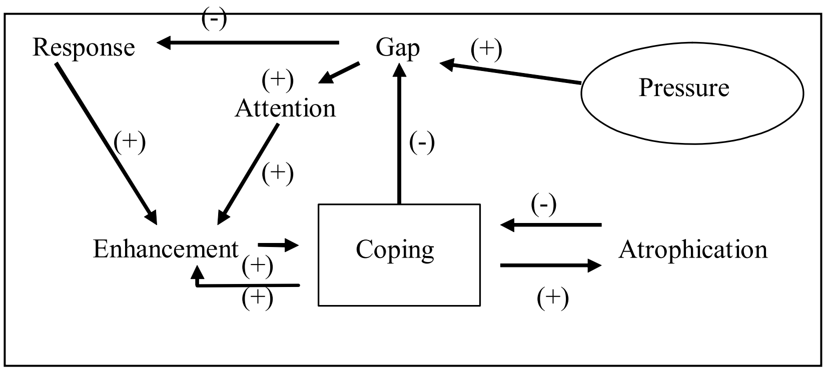


Figure 4: Model Diagram

2.3.1 Basic Runs.

The above equations are now combined together and used to examine coping-skill dynamics and implications. The first simulation shows a simple equilibrium response by initialing coping-skills at unity and holding the pressure at unity. As the results show in Figure 5, the acclimation process reduces the attention the individual gives to the stimulus while simultaneously producing coping-skills improvement. The improvement of coping-skill beyond that needed to accommodate the pressure has both the function of reducing long-term energy needs for coping with the pressure and provides a buffer that allows the individual to accommodate reasonable variations in what may become routine conditions near the new pressure environment. This result indicates that, at the first level of assessment, the model dynamics are consistent with our intuitions and experience.

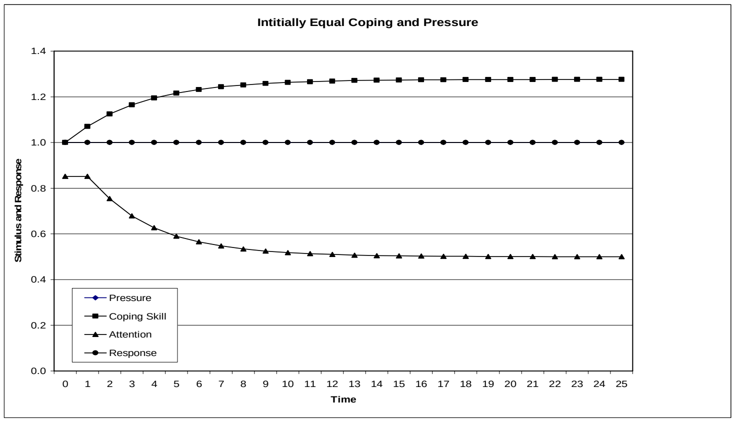


Figure 5: Steady State Pressure.

In all the examples here, the pressure is exogenous. Daily existence is often an unending set of interactions with an inanimate environment and other individuals. Ultimately, the environment or individuals interfere with or constrain the activities of other individuals. This is, by definition, a pressure to which each individual responds. The over compensation of the coping-skill process initiates the equivalent of a coping-skill war that escalates until ether an immovable pressure (an overwhelming constraint) or a pressure relief point (a constraint relief) occurs. An example of an overwhelming constraint is a manufacturing capacity limit or a military response to social turmoil. An example of a constraint relief could be the removal of a rival or the outsourcing of bottleneck-causing problem. Because of this dynamic, most individuals tend to naturally invoke activities that keep them near their coping-skill level, although not necessarily to increase it unless external stimulation requires it. Note that external stimulation that could initiate an increased coping-skill response may be side tracked by another coping-skill response that allows the person to select, for example, quitting an existing job position and finding alternative employment rather than adjust to the new working conditions in the current job. In either case, the individual has selected the path most compatible with existing coping-skill. Entities naturally move to their reference coping levels by "pressuring" the other components of the system or by interfering with others and, thereby causing counter-responses that require the full use of existing coping-skills.

The coping-skill process, thus, has a centering or balancing effect (as discussed earlier) that keeps the individual near the current coping-skill and insures there is no change (that is, no waste of energy) unless an initiating pressure requires it. This feature is inherent to Equation 4. If coping-skill enhancement (CE) of Equation 8 is not big enough to compensate for the atrophication (CA) of Equation 9 (which means the assimilated pressure P is below the coping-skill CS), then CS will decline. Figure 6 shows the dynamics when the pressure falls below the initial coping-skill level and coping-skills atrophy.

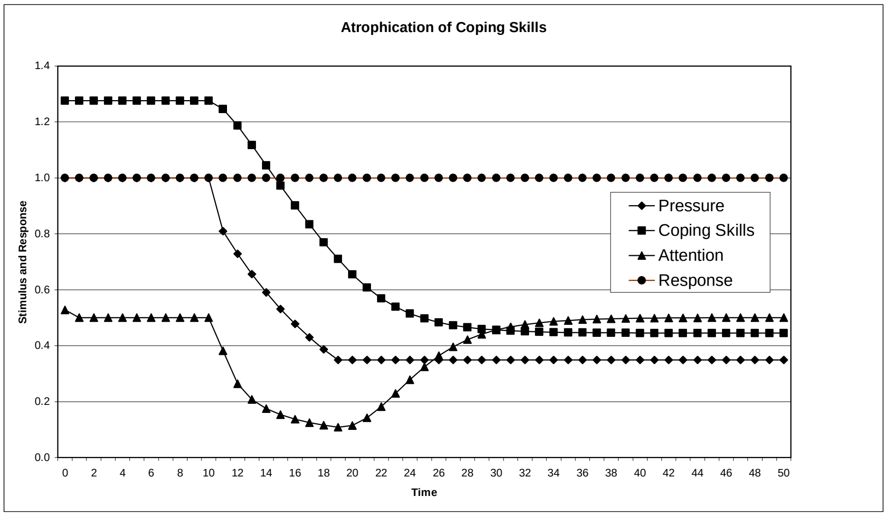


Figure 6: Coping-Skill Atrophication

As the pressure drops, so does the attention. This leads to coping-skill atrophy that falls to match the pressure level. Once the coping skill atrophies to where the pressure and coping-skill are comparable, attention again rises because this level of pressure now produces anxiety and is comparable to the coping-skill level. This result is also consistent with any high-performance activity where the individual slacks off for an extended period.

3.0 Examples of Stress and Coping Dynamics

Because of the shape in Figure 3, there are two equilibrium points for coping-skills. At equilibrium, coping-skill enhancement just matches atrophication. Equilibrium conditions would occur when the pressure and coping-skill place the net response approximately halfway up ***either side*** of the curve of Figure 3. The right side of the curve implies high anxiety and a high energy expenditure. The left side provides just enough simulation to maintain skills, as might be the situation for pilots using flight simulators. The activities of many organizations and governments lead to internal conflict situations that place individuals predominately on the exhaustive right side of the curve.

As an aside, Figure 5 shows the short-term/long-term trade-off of "sitting on one's laurels." In the short-term, anxiety is reduced by avoiding pressure situations, but in the long-term lower pressures produce greater anxiety. This also has relevance to monopolies – political or economic. As noted earlier, once the rival is gone, the pressure is reduced and the coping-skills (the competitive edge) atrophies. The monopolist (or dominant institution) has only a short-term luxury, whether it be VisiCorp (VisiCalc), the Roman Empire, Enron, or Louis XVI of France. Its drop in coping-skill provides the opportunity for another entity to develop coping-skills that surpass it.

Figure 5 may also be illustrative of why assembly line and manufacturing staff seem to have higher stress levels than management. The assembly line flow rate may be completely out of the worker's control. There is no competing against the pressure that the assembly-line brings. Certainly, the survivable response is to give up and let any, now useless, coping-skills designed to control input stimulation to atrophy.

A more severe process is associated with the inability of the worker to be creative or make choices. If there is a rigid, predetermined way to perform a task, coping-skills to deal with variations in equipment performance may be lost. The coping-skills to deal with any change in the routine (routine being the integral concept of coping-skill formation) will have been compromised and any of the bound-to-occur breakdowns or mishaps will represent an overwhelming experience – and a reinforcement to trained helplessness. Many existing and former dictatorial African countries exhibit such a response.

This trained helplessness then makes a change in job routine a threat. Consequently labor-union efforts act in the behalf of employees to prevent job reclassification and task-mobility. Trained helplessness also makes the task of human resource departments all the worse because the coping-skill starting point is so low. If insult is added to injury by having supervisory management who in effect punish line workers or clerical staff for mishaps, then coping-skills and the ability to cope with "routine" problems become minimal. Performance remains at a level far below anyone's expectations, but by necessity, this level of performance must be accepted within the company as the best that circumstances allow.

Despite suffering grueling hours, information overload, and countless job-on-the-line decisions, most managers can make choices. They can often prioritize battles and delegate overflow or responsibilities to others as they see appropriate. They can manage their own risk. They can grow their coping-skills. The stress may be high in absolute levels, but it is typically within a positively reinforcing range relative to coping-skills. For workers without choice options, stresses necessarily become large relative to coping-skills and act to diminished them further. During times of intense economic growth, this phenomenon leads to increasing gaps in performance (and coping-skills) between skilled/professional and unskilled labor.

3.1 Maximum Growth in Coping Skills

There is a maximum sustainable growth rate in coping-skills. How to achieve this rate would be valuable knowledge to employers and psychological counselors. Given the assumptions (parameterization) used here, the maximum growth rate varies around 10% per event-time period. If the time units are months, then Figure 7 indicates that capability can double every 8 months. This is consistent with the requirements of today's competitive high-tech jobs. (Note the change in the Y-axis scale across runs.)

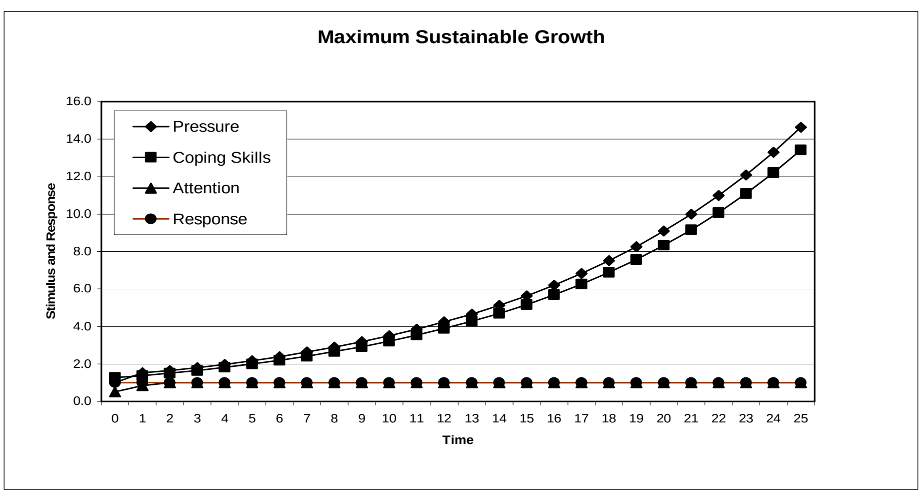


Figure 7: Maximum Sustainable Growth

From Equations 5 through 8, the maximum growth rate occurs when the pressure is CO greater than the coping-skills. Although the syntax used here implies the gap (G) is zero in this case, the gap between the coping-skills and the pressure is at the optimal condition for increased capability (learning).

Figure 8 shows the situation as the pressure exceeds the maximum growth rate. There is still growth, but response is down, as opposed to the situation in Figure 6 where attention and response are both optimized. Comparing Figure 5 to Figure 6 shows that the growth in capability falls far below what could be achieved. The concept of managing individuals by pushing them to their limits as an attempt to maximize societal, economic, corporate, or individual performance appears to be misguided.

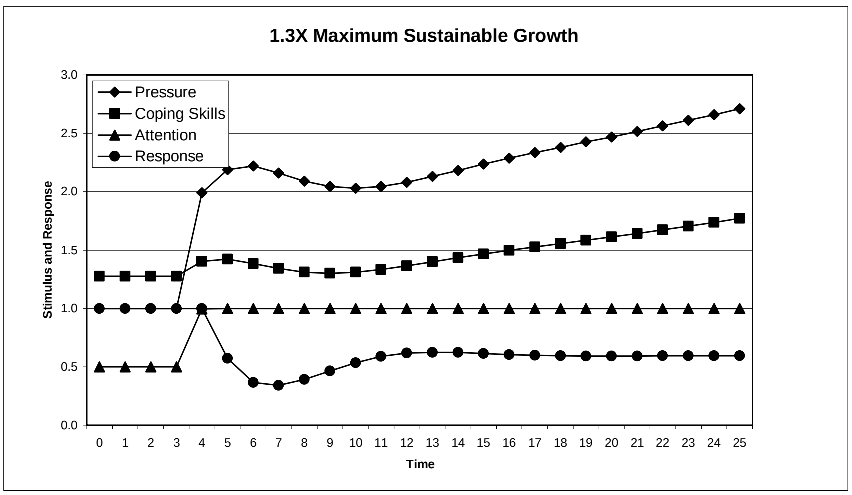


Figure 8: Near the Limits to Growth

Figure 9 shows the response if the pressure is 40% above the maximum sustainable level and thereby exceeds the capability of the entity to respond effectively. The entity responds adequately in the short-term but subsequently becomes overwhelmed and the maximum it can sustain drops precipitously as time goes on. The pressure used in Figure 8 changes continuously to be 1.4 times above the sustainable value, as if a manager continues to find a middle ground pressure between institutional need and expected individual capability. It is a losing situation. Because coping-skill enhancement is a function of coping-skills, once coping-skills drop to low levels, it is a very difficult and long task to bring them back up to proficient levels. Either punishment or excessive demands (pressures) on an individual can quickly lead to a learned helplessness condition.

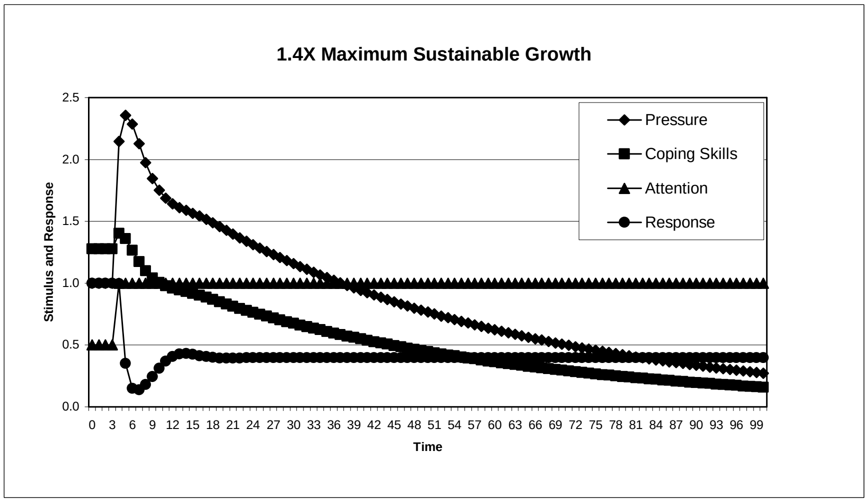


Figure 9: Collapse

4.2 Constantly Escalating Pressure

In most social and business situations, external pressures increase over time. Figure 10 simply increases pressure at a fixed rate of 15%/period. This is 40% to 50% above the sustainable growth rate and corresponds to the pressure that might be experienced in high growth industries or in escalating political/military situations. The individual responds effectively for a while because the pressure starts ***below*** the level that the individual could accommodate. Thus, the early escalation of pressure actually enhances coping-skills and gives the impression of success. Eventually, however, the pressure growth at some point exceeds what can be accommodated and the continued growth soon traumatically overwhelms the individual and response capabilities collapse. Although coping-skills do not fall as quickly as the response, they are far below what is needed to withstand the pressure. In fact, coping-skills peak just before the collapse. Because of this, the individual (or company) could be "revived" by reducing the pressure or by adding the coping-skills from additional staff to the effort.

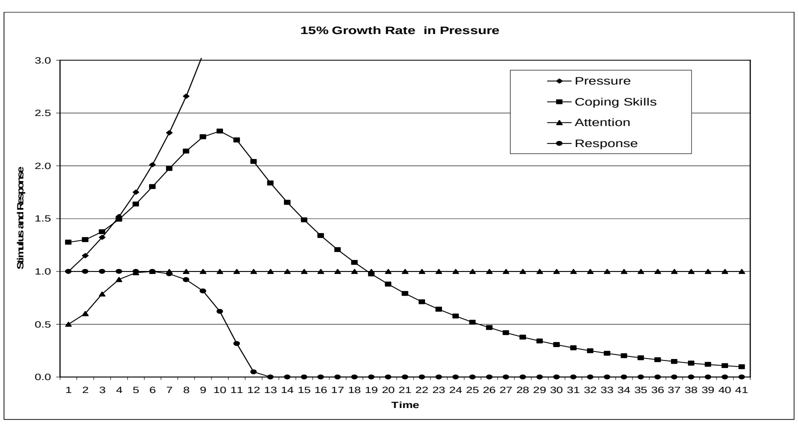


Figure 10: Moderate Coping-Skill Overshoot

Figure 11 shows a situation similar to that in Figure 10 but the growth rate is just above the maximum sustainable long-term rate of pressure growth. Note the qualitative similarity. The rise is higher, the fall further, and the timing elongated. Comparing this simulation to the maximum growth simulation (Figure 7) shows the value of staying near and below the growth limits.

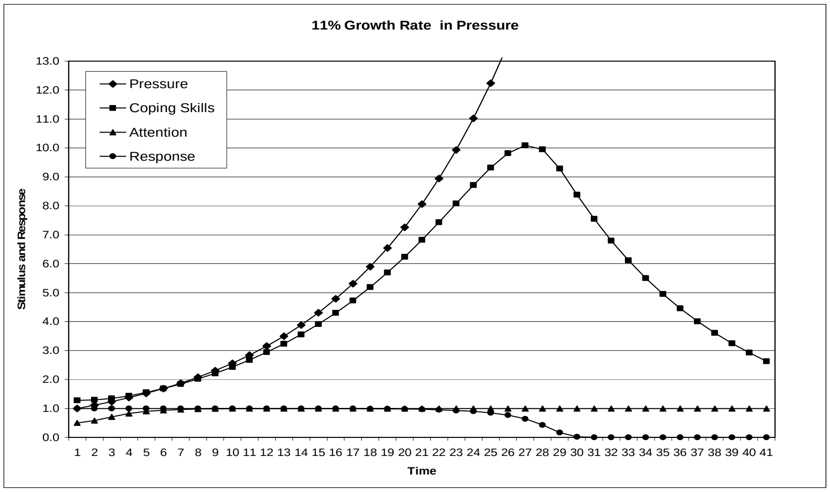


Figure 11: Gradual Coping-Skill Overshoot

4.3 Limits to Coping

Figure 12 shows a simulation that is identical to that of Figure 7 except it includes an upper limit for coping-skills. The individual simply has a maximum potential in the particular coping-skill. Most of us would meet our limits early if we attempted a mid-career change to become either a successful concert violinist or a neurosurgeon.

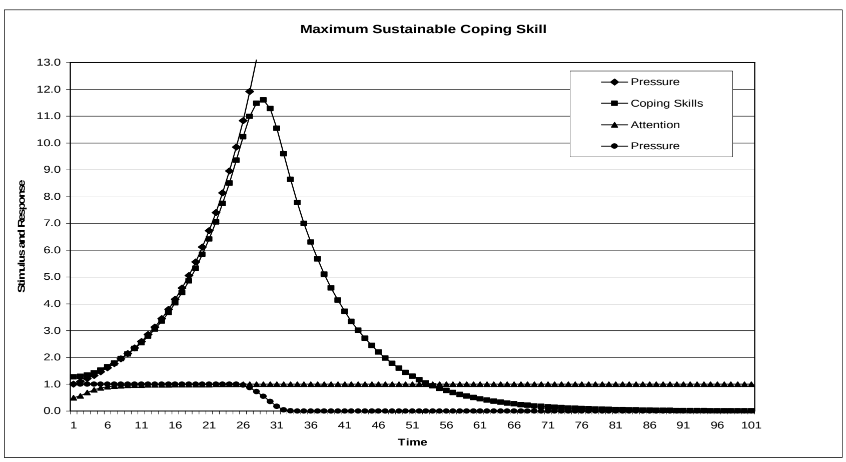


Figure 12: Maximum Sustainable Growth with a Coping-Skill Limit

In the Figure 12 simulation, the maximum coping-skill was set to 10.0 but a 20% short-term boost above that was achieved. This is inherent to the assumption of the curve in Figure 3 and consistent with human capabilities to temporarily perform beyond sustainable levels.

4.4 Repetitive Pressures

Many work-life conditions center on quarterly reporting, sales quotas, and new product release timing. Figures 13 and 14 show the results of simulating that repetitive pressure condition. Both simulations are the same except Figure 13 has a pressure peak 10% higher than that used to generate Figure 14. In Figure 13, eventual burnout occurs. In Figure 14, the individual eventually adjusts to the process. The most significant characteristic is the reverse behaviors of attention and response in the two simulations. In Figure 13, the down periods are a time to rest. In Figure 12, it is just a time not to be as overwhelmed.

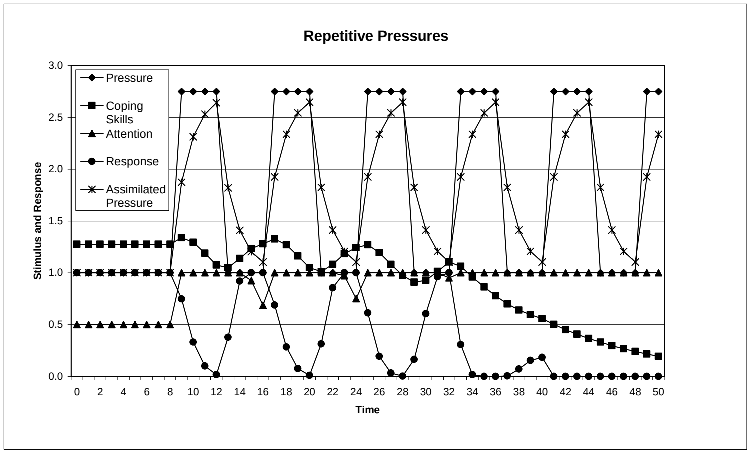


Figure 13: Excess Repetitive Pressure

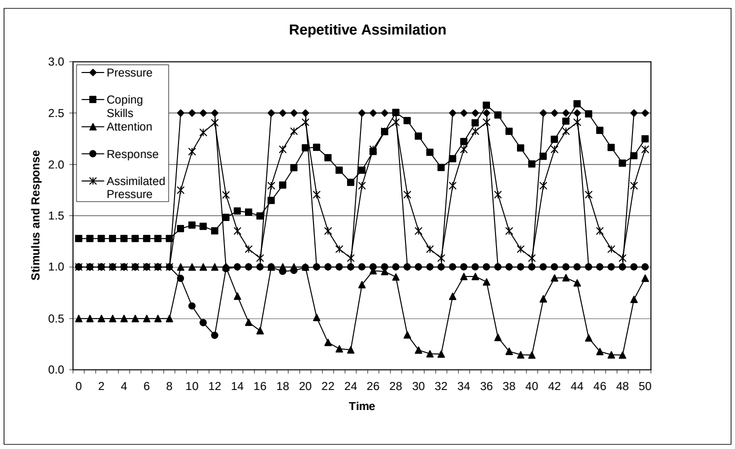


Figure 14: Tolerable Repetitive Pressure

Both Figures 13 and 14 include the additional time series of Assimilated Pressure (Equation 11). Because an individual cannot immediately recognize the extent or the duration of a pressure event, it takes time to assimilate the pressure. The assimilation is a delay process and attenuates the experienced pressure. During the early stages of the excessive pressure increase, coping-skills also increase, provided the assimilated pressure remains within manageable bounds. This dynamic process allows an individual to "rise to the occasion" (to some extent) and accommodate what would otherwise be an overwhelming pressure. While this dynamic appear to be mostly applicable to the quarterly burn-out of accounting departments, it may more help explain long-term anthropological processes. Suppose the variation is on different time scale, such as seasonal water flows or growing periods. Those who survive the most hardships should eventually be the most advanced. Much of North America is either dry or cold. Yet, it is in South and Mesoamerica where the ancient civilizations flourished. If these peoples initially came from the North (East Siberia), coping theory suggests that we would expect to see the equivalent of the Aztecs in Oregon. The southern location of these civilizations indicates southern (SE Asian/Polynesian) origins. And, indeed, recent evidence continues to support that position compared to the conventional Clovis origins (Dillehay, 2003).

On the reverse side, it takes a while before an individual realizes a reduction in pressure. The pressure assimilation process allows the past to linger – just in case it comes back. This was probably a very useful quality when humans were not at the top of the food chain. Most importantly, it temporarily maintains coping-skill growth after the removal of the pressure. Even if there is a slight decline over the peak coping-skills, the overall coping-skill could remain high because the individual will soon interact with the environment in way that again increases the pressure. Hence, limited-duration relaxation after coping-skill overload events lead to future productive response capabilities.

4.5 Burn-Out

Burn-out is phenomenon commonly applied to and commonly recognized in individuals who spend time in a environments of enduring, high stress. In the simulation of Figure 15, a pressure just below that used for Figure 13 is used and sustained. Even with the delayed recognition of assimilated pressure, the individual is almost overwhelmed and falters, but eventually recovers and develops adequate coping-skills for the new environment.

Lastly, Figure 16 increases the pressure just to the point where it is unsustainable – just a tad higher than that used in Figure 15. As shown in the pressure growth simulations, the ramp-up time for the pressure is more crucial than the final magnitude of the pressure. The ramp-up rate and magnitude is so high in the Figure 16 simulation that the individual just begins to respond before becoming overwhelmed. The coping-skills are not affected immediately because the inability to respond dominates. This simulation represents the classic “burn-out” phenomena. A removal of the pressure could leave the individual with greater coping-skills than he/she had initially.

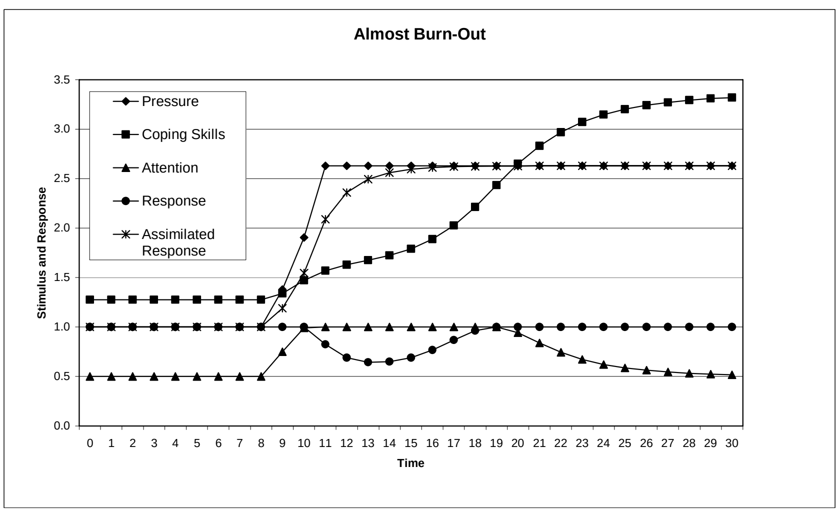


Figure 15: Almost Burnout

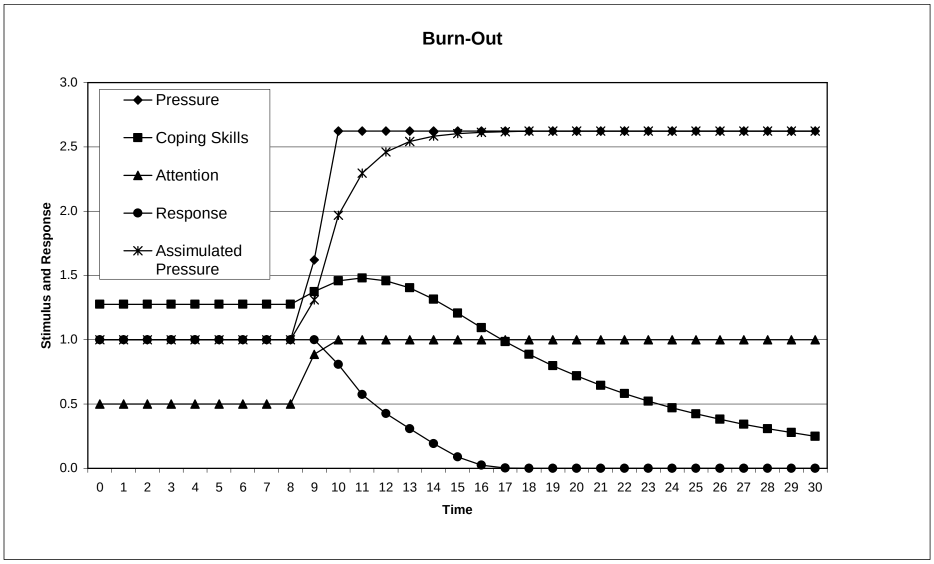


Figure 16: Burnout

4.0 Discussion

The simulations and associated representation of coping-skills presented here have wide applicability. The original research of others, as noted earlier, helps explain and treat alcoholism and worker burnout. As a generalization of alcoholism, drug abuse both reduces coping-skill and acts as a means to avoid dealing directly with the environment that could improve coping-skills. It becomes a self-reinforcing means of coping unto itself that prevents non-drug-augmented, that is, real, coping-skills from establishing themselves. The consequential atrophication of other coping skills generates a learned helplessness to perform adequately under normal condition. Criminal delinquency also has this feature. This would indicate that rehabilitation requires a "starting over" that recognizes that "normal-coping skills" have almost vanished and initial progress, to be solid, needs to be slow.

As intimated in the earlier discussion of bullies or rogue national leaders, violence is often a response that stems from incapability rather than capability. The same logic would seem to apply to racism and all forms of prejudice. This work would indicate that as nations become more capable and self-sufficient, they will feel less threatened and the risk of major turmoil should diminish. This assertion does imply a democracy, however, because all dictatorships are like the monopoly noted earlier that stimulates external change – that with which a dictator cannot cope.

In a reverse sense, it is crucial for all nations, people, and societies to feel a certain sense of inadequacy under a burden of pressures. Otherwise, their coping-skills would atrophy and the negative dynamics of inadequate coping-skills would ensue. These pressures do represent conflicts, but when all parties have adequate coping-skills, conflict resolution focuses more on the resolution than the conflict because the natural tendency is to have the two positions balance. That is, in a two-way conflict, each party is the pressure to the other, as precipitated by the respective coping-skill. The two pressures will equilibrate, as will the coping-skills.

Just as prices imply the equilibrating effects of the invisible hand in economics, coping skills imply the invisible hand in the social context. In fact, a successful economic market is simply a collection of interacting coping-skills negotiating a price. Those who cannot cope lose, and those who have appropriate coping-skills (or luck) succeed.

Fortunately, individuals and societies have many pressures to increase their skills. People and societies migrate, use common resources, and expand geographically to eventually interfere with each other and to produce pressures. The human-centered world became increasing interconnected from 9000 BC onward with subsequent interactive pressures that led to the early civilizations and related conflicts.

"Change management" and "learning environments" have become key terms in corporations. Change and learning are the basic features of coping-skills. The problem is that there are lots of coping-skills, and "change management" and "learning environments" often have general definitions that misguide the effort to less than hoped for consequences. This work would indicate the need to delineate what coping-skills define the goal and what are the limits of the growth path to achieve the goal? Then the task becomes what pressures to ***impose*** on the group or organization and at what rate to improve those skills. Learning and change can only occur in response to a pressure demanding them. The effort must be to improve the existing skills rather than create them, because additional coping-skills can only build upon existing coping-skills. If the skills do not exist, then the questions become 1) have the wrong skills been recommended for change, or 2) can the existing skills be enhanced to eventually perform the tasks needed?

The coping-skill theory presented here would argue that individuals naturally push themselves in the areas where they succeed and avoid those areas of their weaknesses. This comes about 1) because successful coping-skill responses grow beyond that needed to counter a pressure so as to allow efficient response for all variations of that pressure, and 2) because the coping-skill itself is a pressure to some other entity in the system that then induces a pressure to reinforce an already successful skill. Thus, a theory of coping-skills helps explain motivation. The needs most associated with motivation are physiological needs, material comfort, security, and freedom (Christensen & Norgard, 1974; Maslow, 1970). If freedom is deemed the ability to choose and security the ability to cope with external forces, then the concept of coping-skills would seem applicable.

The fact that many historical figures have flourished under incredible odds supports the view that coping-skills, in general, have no limits even though individual coping-skills do. The coping-skill description proposed here would imply that many of these handicapped individuals rose to greatness because of their handicap rather than in spite of it. The plights and successes of Helen Keller and Stephen Hawking are common knowledge. Winston Churchill, Leonardo Da Vinci, Albert Einstein, Thomas Edison, Alexander Graham Bell, Werner Von Braun, H.G. Wells, George Washington, General George Patton, and Woodrow Wilson all had severe dyslexia. Alexander the Great, Julius Caesar, Alfred the Great, Napoleon, Charles Dickens, and Van Gogh were epileptics. They all overcompensated for limitations often in the specific area where the limitation was most crucial. To perform complex mathematic derivations without a physical memory/organization aid, as in the case of Stephen Hawking, is the last thing most physicists or mathematicians would like to imagine. Planning for battle is complex enough with out having to include a contingency for an epileptic fit as in the instance of Alexander and Napoleon. The dyslexia of Einstein and Edison forced them to view the world differently to great ends. Early success at compensating for the handicap and the continued interference of the handicap caused coping skills that allowed achievements that would normally be deemed impossible.

In a related vein, this view of behavior then casts the role of a counselor or therapist as a force for reorganizing existing coping-skills to a compatible set and finding those functional skills that can be used to compensate for those that are severely underdeveloped or have limiting maximum levels. The role of the "teacher" is then to increase the desired (needed) coping-skills at a rate consistent with their existing level. In this context, Figures 6 through 9 can be viewed as illustrating a problem with conventional schooling. Ideally, every student would have the response of Figure 7 with maximal growth. Because of the varying level of coping-skills of individuals in the classroom, the teacher must opt for an "average-student" approach and achieve the scenario in Figure 8 where growth is low for all and stress is high for many. Unless the child's coping-skills are mature enough to cause the child to continuously generate external or internal pressures that increase coping-skill development, the limiting atrophication of Figure 9 occurs, albeit followed by the relatively weaker growth of Figure 8, but in an under-stimulated sense.

Many adults clearly have coping-skills thousands of times in excess of those of a young child. Given that the implied growth rates literally reflect a building of coping-skills upon coping-skills, those children who do not receive a strong early boost in coping-skills have a much lower chance at long-term success.

A worse situation, as illustrated in Figure 7, is for those students who have been left behind. These are the students whose level of abilities (coping-skills) were no match for the initial level of stimulus (pressure) and are certainly incapable of accommodating an upward growth from that point. Not only do they spiral down, but also the gap between the successes of those students on the trajectories of Figures 7 and 8 and the failures of those on the trajectory of Figure 6 grows indefinitely, possibly throughout life. This bifurcation into the "good" versus "bad" trajectory is abrupt and the separation grows quickly because of the rapid positive reinforcement of both the "good" to be better and the "poor" to grow worse. Through this process, societies and nations separate into ever widening divisions of "haves" and "have-nots."

For example, this logic can unfortunately be expanded to compare the indigenous societies of Africa and North America. The steps "forward" in the colonial days of Africa were not only problematic because of the exploitation, but because they caused changes that were too dramatic to ever be accommodated. The coping-skill dynamics parallel the discussion of the overwhelmed worker associated with the simulation of Figure 6 above. Learned helplessness made the indigenous people worse off than they initially were. Nonetheless, the key point is that loss of capability would have happened whenever a group faces off against the overpowering capability of another. Certainly there are individuals within the group who could either compensate or directly rise to the task, but the average response conforms to the logic inherent to Figure 6.

The fact that indigenous peoples could not change quickly enough to meet the labor-performance expectations of the usurper society, caused increased tensions that often led to punishment as an "incentive" to improve performance. The relative limitations of the indigenous peoples compared to the occupiers, then brought about slavery and institutionalized learned helplessness. This is quite different from the slavery of the Greek or Roman times, where the coping-skill (capability) gap between victor and victim was either nonexistent or not so large. It is possible that the existing gap between the performance of different ethnic and racial groups in America reflects the perpetuated gap in coping-skills developed in those earlier times. This logic would also then apply to long-term economically disadvantaged individuals existing in an affluent society.

It may be that in schools, third-world development, and welfare programs, a big step backward to develop rudimentary critical coping-skills is needed before conventional "forward" steps can succeed.

4.0 Summary

In summary, an individual's set of coping-skills defines his/her personal existence, as well as that of our society and all subgroups the compose society, over all times. Coping-skills may be the defining approach to understand both the human condition and the evolving environment in which it competes. The stresses in our lives are what make our lives. Both too much stress and too little stress has negative effects. Despite the fact that much of the philosophy of Nietzsche is distasteful, he appears to have been correct when he said, "That which does not kill you, makes you strong."

Acknowledgment:
The author is grateful to Ralph L. Levine's expertise, comments, and encouragement for this study.

R**eferences:**

Atkinson, J. W. and N. T. Feather (1966). *The Theory of Achievement Motivation*. New York:Wiley

Atkinson,. J. W. (1964). *An Introduction to Motivation*. Princeton: D.Van Nostrand

Backus, G. (1980a). *Integrated Energy Analysis.* (Dissertation) Resource Policy Center, Dartmouth College, Hanover, New Hampshire, DSD 194 Volume. 1, pp. 28-26

Backus, George A. (1980b). *Human Behavior: Culture vs. Biology*. Report No. AFC-1, School of Industrial Engineering, Purdue University, West Lafayette, Indiana

Ben-Akiva, Moshe and Steven Lerman (1985) *Discrete Choice Analysis*, MIT Press, Cambridge, Massachusetts. Pg. 360.

Bock, R. and L. Jones, (1968). *The Measurement and Prediction of Judgment and Choice*, San Francisco: Hoden-Day

Christensen, B.L. and Norgard, J.S., (1974). *Social Values and the Limits to Growth*. Thayer Scholl of Engineering, Dartmouth College, Hanover, New Hampshire, DSD 17

Dillehay, T.D. (2003). Tracking the first Americans. Science 42, 23-65

Donnadieu, Gerad, and Michel Karsky (1990). The Dynamics of Behavior and Motivation. In *Proceedings of the 1990 International System Dynamics Conference*, July 10-13, 1990, Chestnut Hill, Massachusetts (pp. 319-332), Albany: System Dynamics Society.

Forrester, Jay, (1961). *Industrial Dynamics*: Appendix E. Cambridge: MIT Press,

Goebel T., M. R. Waters, and M, Dikova (2003). The Archaeology of Ushki Lake, Kamchatka, and the Pleistocene Peopling of the Americas., Sciecne 301, 501-505

Golüke, U. (1980). *A Comprehensive Theory of the Pathogenesis of Alcoholism.* Dartmouth System Dynamics Group, Dartmouth College, Hanover, New Hampshire, Report DSD #252

Golüke, U. (1983). The Pathogenesis of Alcoholism. In B. Kissin and H. Begleiter (Eds.), *The Biology of Alcoholism***,** New York, NY: Plenum Press, pp. 605 - 675

Golüke, U., Robert Landeen & Dennis Meadows, (1984). A Simulation Model of Drinking. *British Journal of Addiction*, 76, 289-298

Helson, H. (1964) *Adaptation-level theory; an experimental and systematic approach to behavior*. New York: Harper & Row

Homer, Jack, (1985). Worker burnout: a dynamic model with implication for prevention and control. *System Dynamics Review*, John Wiley, Vol. 1, Number 1, 42-62.

Hunter, J. E. , J. E. Danes & S.H. Cohen (1984). *Mathematical Models of Attitude Change: Volume* . New York: Academic Press

Lazarus, R. S. and S. Folkman (1984). *Stress, Appraisal and Coping*. New York: Springer

Levine, R L., Mary Van Sell, and Beth Rubin (1985). A Model of Burn-Out. In The Workplace. *Proceedings of the 1985 International System Dynamics Conference*, Keystone, Colorado, July 2-5, 1985. (pp. 487-501). Albany: System Dynamics Society.

Levine, R. L. (2003). Models of Attitude and Belief Change from the Perspective of System Dynamics. In *Proceedings of the 2003 International System Dynamics Conference*, July 20-24, 2003, New York (page 84: electronic link), Albany: System Dynamics Society.

Luce, R.D. (1959). *Individual Choice Behavior: a Theoretical Analysis.* New York: Wiley

Luce, R. and P. Suppes (1965). Preference, Utility and Subjective Probability. In R. D. Luce, R. R. Bush & E. Galanter (Eds.), *Handbook of Mathematical Psychology*. Vol. 3.. New York: Wiley

Maslow, Abraham H. (1970). *Motivation and Personality* New York: Harper and Row

McFadden, D., (1986). Econometric Model of Probabilistic Choice. In C.F. Manski and D. McFadden (Eds.) *Structural Analysis of Discrete data with Econometric Applications*, (pg 198), Cambridge: MIT Press.

Moos, R. H. (2002) The Mystery of Human Context and Coping: An Unraveling of Clues. *American Journal of Community Psychology*, Vol. 30, No. 1, pp. 67-88.

Popper, Karl R. (1982). *Quantum Theory and the Schism in Physics* (pp. 1-34), New Jersey: Rowman and Littlefield

Simon, H. A. (1957). *Models of Man*. New York: Wiley & Sons

Wolfram, Stephen (2002) *A New Kind of Science*. Champaign, IL: Wolfram Media, Inc.

Appendix 6: A Network ABM Approach to Self-Organized Extremist Group Formation, Activation and Dissipation

Written by RJ Glass, 3/14/05
Revised with WE Beyeler, GE Barr, 6/6/05
Revised 10/28/2005 after meeting with Colbaugh, Backus, Beyeler, LaViolette, Oishi
Edited by George Backus 1/20/2006

This section presents a simple networked agent-based model that abstracts the essence of self-organized extremist group formation, activation, and dissipation in context of a structured society (multiply overlapping groups) and where the influence of external and internal drives and perturbations can be evaluated. It can be used to identify system features, attributes, types of perturbations and drivers that result in extremism and determines strategies that can reduce or enhance the likelihood of formation, activation and/or dissipation of extremism.

General Approach Model Description:

The general model is a network based extension of the work of Epstein (2002) that represents/generates the psychological/sociological processes identified within the ABM sketch through application of fundamental behavioral mechanisms (Sageman 2004). Within this construct, social networks are composed of multiple overlapping groups (work, school, sports, gangs, etc.) based on the demographics of the society surrounding the group of interest. Individual agent attributes (Status, Plasticity, Tolerance, Risk Aversion) and initial values of at least one "state" or "belief" variable (such as Hardship and/or Legitimacy of local regime, leadership or societal norms) may be correlated to class, group, position within group, local network topological measures (degree, clustering coefficient), or global topological measures (critical node analysis). In addition, an agent's position in one group may be correlated to their position in another (i.e., CEOs tend to be on the boards of other CEO's companies; group leaders within the community tend to be members of the Rotary Club).

From an initial distribution and structure within the social network, belief levels change via 1) individual to individual contacts, 2) group participation contacts, and 3) secondary impersonal contacts from literature, the media and/or the internet. Individual to individual contacts may be thought of as a game where the outcome is either synergy, anti-synergy or compromise. In this game, the distance between the two individual's entering belief levels (Delta), the individual's attributes (Status, Plasticity, Tolerance) and some degree of randomness (Noise) are incorporated to modify belief levels of the agents. There may also be preference falsification and lying. The outcome of the game also adjusts the frequency of link activation between the two agents, or strengthens, weakens, or entirely severs their link. Group participation contacts and secondary impersonal contacts provide multi-scale feedback to the individual's beliefs, which could have comparable outcomes to those for individual to individual contacts, and could involve the same or additional inputs to determine final belief levels. Belief levels influence individual strengths of conviction for action, i.e., Motivation. In context of Epstein (2002), Motivation is synonymous with his "Grievance" (or as earlier defined as “dissonance”). Individual action states (i.e., vocal, seeking, joining, training, plotting, inciting, and ultimately participating in violence) are defined by Thresholds in motivation. These action states 1) provide a visible queue to the agent's belief level that other agents can see and be influenced by via herding, and 2) cause additional behaviors to be implemented that modify the agent's social network (i.e., searching for like minded agents, forming new bonds, breaking old bonds). Agent attributes tied to the position of the agent within the various groups (Status, Plasticity, Tolerance) will also change as the agent modifies his networks. Other agent attributes such as Risk Aversion could also change or be made to be a function of the time that an agent remains in a state.

Action state thresholds are also influenced by a herding process through evaluation of Net Risk weighted by the agent's Risk Aversion. Net Risk includes that of being chastised by those whose opinions the agent values, being reported by others (an increasing function of distance outward within the social net from 0 to some asymptotic average value), or being arrested by the government. The consequences for each of these would be different and would need to be imposed. For example, Epstein (2002) removed agents that were arrested and put them in jail for different lengths of time. While in jail, beliefs and connections to others could be modified in a variety of ways to consider various policies.

Initial implementation for two-poled extremism:

Initial work uses the opinion-adjustment logic noted here rather than the more complex logic noted earlier. The logic below can still function with the limited data expected during the initial phase of the M&I program.

Based on Epstein (2002), each agent has two state or belief variables: Perceived Hardship, H (combination of economic, opportunity, alienation...), and Perceived Legitimacy, L, of the status quo (range from apologists to provocateurs). SNL research extends Epstein's (2002) single pole model to two poles of extremism: P**rovocateurs** (proponents of change) and **Apologists** (proponents of the status quo). These provocateurs add the “break” dynamics that the original model misses. An agent's perceived **Legitimacy** goes from -1 (extremist for status quo, apologist) through 0 (neutral) to +1 (extremist for change, provocateur). An agent's perceived **Hardship** goes from 0 (none) to 1 (extreme) Discrete action states are defined by **Thresholds** (one for each edge of the action state) within a **Motivation** spectrum with:

Motivation = Hardship*Legitimacy

An agent falls within an action state if his motivation modified by his perceived **Net Risk** of being in that state is above and below the two thresholds defining the action state.

The structure of Legitimacy evolves within the social network with individual to individual contact, as imposed by the social net (links, frequency of link activation) The Legitimacy levels after interaction should be dependent on the difference between the two agent's Legitimacy, |deltaL| with our initial thinking:

-if |deltaL| > deltaNeg: anti-synergistic, both agents move apart

-if |deltaL| < deltaPos: synergistic, both agents move toward the common pole

-if deltaPos < |deltaL| < deltaNeg: compromise, both agents move together

However, in the simple model currently implemented, we ignored the synergistic interaction and tied everything to a tolerance:

-if |deltaL| > scaling2antagonize*Tolerance: antagonize, both agents move apart

-if |deltaL| < Tolerance: compromise, both agents move together

-if Tolerance < |deltaL| < scaling2antagonize*Tolerance: ignore

Note that for the synergistic and anti synergistic interactions the quantity "|L agent 1| + |L agent 2|" is not conserved, while for the compromise, it is. Hence, polarization will always result with this rule set. The outcome of the interaction is a function of the Status of the two agents and their Plasticity or ease to which they can be influenced by others. Both Status and Plasticity can be tied to position within the local group hierarchy; it could also be related to the degree of the node. There is a need to model an information filter that includes misunderstanding (stochastic component to the game outcome with tunable Noise) and falsification where agents can only see the external queue given by the agent's state (and thus L falsified due to Risk Aversion). Trust, possibly based on repeated interaction wears down Risk Aversion of the agent and allows the true L of the agent to be recognized by particular interacting agents. As a result of the game outcome, agents also adjust their frequency of link activation and possibly strengthen or weaken their link. Links can be severed completely when they fall below a given frequency.

There is a "group meeting" influence that is different from individual-individual contact. Groups create environments for strong bonding with an alignment of L and the possible synergistic shift of L towards a pole. An agent's Legitimacy level is also influenced through larger scale "interactions" or "proclamations" from leaders, or sub-leaders, or sub-sub-leaders of groups and from media (radio, TV, or same on internet).

Action states encompass at least the following:

0) Passive
1) Vocal
2) Seek and Join extremist groups, Decrease frequency or Sever old links
3) Train and Plot
4) Incite or Participate in Violence (Apologists can turn to vigilantes and Provocateurs to terrorists)

Individual agent action state transition:

A simple first order model for action state transitions by an individual agent can be fashioned as an extension of Epstein (2002). Consider each agent to have a level of Motivation, M, that is a function of the two state variables H and L:

$$M = HL$$

At a minimum, two sets of thresholds are defined for the population - one for Provocateurs, Tp_i, when Motivation moves toward the positive pole, and another for Apologists, Ta_i, when Motivation moves toward the negative pole. (Note, these thresholds could be agent class specific, or even variable from agent to agent). Each set contains the 3 (or more) action level states (denoted by subscript i) presented above. For example, if an agent's Motivation is pushed above Tp_1 or below Ta_1, they may move from the inactive action state Tp_0 or Ta_0 to either the level 1 Provocateur or Apologist states, respectively. Whether an agent transitions to another state is dependent on the agent's evaluation of Net Risk, Np_i or Na_i, for taking on the behavioral pattern of the i^{th} state. Combining Net Risk and Motivation and comparing to the specific thresholds determines the action state of the agent:

$$Tp_{i+1}>M-Np_i>Tp_i \text{ or } Ta_{i+1}<M+Na_i<Ta_i$$

Where N is a product of an agent's level of risk aversion, R, and the probability, P, of being chastised, reported, or imprisoned/killed for their actions by other known agents of the opposing view:

$$Np_i=RPp_i \text{ or } Na_i=RPa_i$$

As an initial step, we follow Epstein in taking P as an exponential distribution dependent on the number and state of Provocateurs and Apologists that the agent is likely to have contact with within a proscribed period of time, t:

$$Pp_i=1-\exp[-k(\text{ApologistsAboveState0}/\text{ProvocateursAboveCurrentState})_t]$$

$$Pa_i=1-\exp[-k(\text{ProvocateursAboveState0}/\text{ApologistsAboveCurrentState})_t]$$

Where k is set to yield a plausible estimate when the number of active Provocateurs and Apologists are equal (P of 0.9 is what Epstein used for this situation to determine k).

Agents are introduced to other agents where they may form new links. A frequency of "introduction" could be associated with each contact whereby an agent introduces one of his links to another which have not previously been linked. An introduction interaction plays the same individual to individual interaction game; the outcome of which not only influences the Legitimacy levels, but whether a link is formed, the frequency of activation, and strength. Agents are offered random individual to individual interactions in time. There is the directed search to find others of like mind: Provocateurs and Apologists actively search out like when in level 2 or above. The risk of going outside of their social net can be modeled as an increasing function of distance outward within the social net from 0 to some asymptotic average value. The search could be to query their own social net first (lowest risk), then the social nets of their friends (friend of a friend, moderate risk), then a random walk to find other level 1s and above (high risk). Here, the model could also consider the internet mechanisms (vocal, chat, email) for introduction (highest risk).

The frequency and strength of links are adjusted and links possibly severed due to the interpersonal interaction game. New links are added over time as outlined above. There is a limit to the number of substantial contacts that an agent can have in a given length of time. Therefore, link frequencies must be adjusted to keep an agent in bounds; low frequency or low strength links must be removed. Also, when an agent finds and joins an extremist group (level 2), they actively drop other links as they take on new extremist group links.