**Project Title:** Strength and Corrosion in Social Norms: Conditions for Norm Adherence

**Student Researcher:** Alexander Leeds

**Primary Faculty:** Rachel T.A. Croson
Maurice E. Schweitzer

**Descriptive Summary:**

Some of the most important questions in the area of decision making under uncertainty deal with individual decisions to conform to or violate norms of behavior. The decision to conform to a norm is complicated by uncertainty about the norms others hold (Bohnet and Zeckhauser, 2004) and strong inclinations to behave collectively (Milgram, 1977). (For the present purpose, we adopt the broad definition of norms as shared standards of behavior (Thibaut and Kelly, 1959)).

This project proposal is part of dissertation research into the generalized conditions that lead to adherence to norms across different norms and contexts. This is a crucial matter because several challenging social concerns, not least matters of crime, consumption of natural resources, and interdependent security, may partly be resolved by promoting conformity to deliberately chosen norms. Identifying and comparing the motivations for adherence to norms is consequently an open and growing focus of research (Davis and Holt, 1993 (p. 332); Elster, 1989; Hernandez and Curhan, 2005; Ostrom, 2000).

We advance emergent norm theory (Killian and Turner, 1957) by building on the large body of experimental economics findings that describe factors that independently promote adherence to norms (see, for example, Axelrod, 1984; Deutch, 1958, Friedman, 1967; Ledyard, 1995). However, we examine how these factors behave in concert, rather than independently, to support a generalized and mutable norm construct. Toward this purpose, we examine norms in three settings: the dictator game, the minimum-effort (coordination game), and the standard public goods (prisoner’s dilemma) game.

We isolate three of the most commonly cited factors that strengthen adherence to norms:
1. The existence of a mechanism to establish a precedent for norm conformity (including consensus achieved through communication)

2. Common knowledge that norm adherence improves group welfare over the outcome of individual self-interested behavior (including knowledge of the benefits of norms as group identifiers)

3. The ability for participants to punish others for nonconformity, this last implying some level of visibility of participants to each other

These factors are often found in combination in naturalistic settings where norms have been successfully created and sustained (e.g. “Buy Japanese” in Japan (Elster, 1989), norms for water allocation in India (Bardhan, 1999), norms for groundwater allocation in Southern California (Blomquist, 1994), and reciprocity in workplace contribution (Chatman and Barsade, 1995; Mannix, Neale, and Norcraft, 1995)). Yet, the most common finding in the laboratory is not the survival of the norm. Instead, a regular decline in adherence to norms is usually observed as participants “learn” rational behavior across multiple of rounds trust, public goods, or common pool resource games. (Andreoni, 1988; Davis and Holt, 1993; Engle-Warnick and Slonim, 2004). We aim to resolve this discrepancy between norms’ survival in naturalistic settings and norms’ corrosion in the laboratory by examining how precedent, belief, and punitive tactics function in combination to increase individual tendencies to conform to norms.

Our experimental design makes conservative use of our intended 150 participants. Every participant therefore plays two games from amongst the three possible games, the dictator, minimum effort, and public goods games, in separate 30 minute sessions. The same pairs of semi-anonymous or anonymous (depending on treatment) participants play five practice rounds and ten actual rounds of a game before being randomly paired with a new partner for a new game under new conditions. Within each game, we have four conditions. In the control condition, all naturalistic elements hold for both members of a pair: precedence, common belief in welfare benefits, and the availability of punitive tactics. In the three other conditions, one of the three factors is removed while the remaining two hold. Therefore, 12-13 unique pairs complete each of the 12 possible combinations of games and conditions. We anticipate complete conformity to the norm only in the control condition, and reversion to previously observed behavior when any element of the control condition is relaxed. We have no a priori hypothesis about which experimental condition will result in the most rapid corrosion of the established norm.
Research Expenses:

- Regular Payment for Behavioral Lab Participants: $450
  \(30\% \times 150 \text{ participants} \times \$10\)
- Variable Payment for Behavioral Lab Participants: $750
  \(150 \text{ participants} \times \$5\)
- Research Assistance: $600
- Conference Travel (IACM, 2006): $800
- Computer Programming Expenses: $1500
  (for services not supplied by OPIM Dept.)
- Materials (including photocopying): $200

**Total:** $3,500

There are no additional current sources of research funding.

References:


Gardener, and James Walker, eds., *Rules, Games, and Common-Pool Resources*.
Ann Arbor, MI: The University of Michigan Press.


Selection Mechanisms and Reputation Formation

Jennifer Dunn

Ph.D. student, Operations and Information Management Department

8-0553 / jrdunn@wharton.upenn.edu

Project Description

Trust is an important component of many economic and social decisions (Dirks & Ferrin, 2000). Oftentimes, individuals are faced with trust decisions about trustees with whom they have had little personal experience. In these cases, decision-makers often rely on the reputation of the trustee when making their decision (Dasgupta, 1988). Prior research has demonstrated that a positive reputation leads to more opportunities in social exchange (Dasgupta, 1988) as well as advancement in an organization (Ferris, Blass, Douglas, Kolodinsky & Treadway, 2003). As new organizational forms emerge (e.g., strategic alliances and virtual organizations), personal reputations will become increasingly important across organizational boundaries (Sheppard & Tuchinsky, 1996). Given the value in possessing a favorable reputation, the mechanics of individual reputation formation is an important and understudied question. In particular, we know very little about what types of reputations (e.g., reputations for intelligence vs. reputations for integrity) propagate in organizations and what drives this propagation.

Formal models of reputation formation have generally taken one of two forms: atomistic or fully embedded (Anderhub, Engelman & Guth, 2002). In an atomistic model, a person’s behavior towards a counterpart affects his reputation with that counterpart, but only with that counterpart. In a fully embedded model, a person’s behavior towards a counterpart is observed by, or communicated to, all future counterparts, such that everyone has full knowledge of the person’s past history. Outside of these formal models,
researchers have recognized that communication plays an important role in reputation formation (Burt & Knez, 1995; Uzzi, 1998), but few researchers have examined how people decide to initiate communication about others’ traits and behaviors. The scarcity of research on the role of communication in reputation formation is surprising given the ubiquity of gossip in organizations (Noon & Delbridge, 1993), our propensity to share certain types of experiences and information (Heath, 1996; Rime, Mesquita, Philippot, & Boca, 1991), and the significant effect that communication has on impressions (e.g., Richins, 1983).

In this thesis, I hypothesize that people will communicate about some attributes/behaviors more often than others and that individuals possessing oft-communicated attributes will develop more widespread reputations than individuals with seldom-communicated attributes (e.g., dishonest individuals’ reputations for dishonesty will be more widespread than honest individuals’ reputations for honesty). I hypothesize that individuals are more likely to communicate about behaviors that (1) are strongly rewarded or punished, (2) are perceived to be diagnostic of the person’s traits, and (3) generate emotion in the communicator. I draw from prior research in these areas to hypothesize specific patterns of reputation formation. For example, I hypothesize that favorable reputations for creativity will be more widespread than unfavorable ones, but that unfavorable reputations for integrity will be more widespread than favorable ones.

Methodology

In this proposal, I plan to test my hypotheses using a longitudinal study of reputation formation within two to three student organizations. Specifically, I plan to create a website through which participants periodically evaluate the other members of their organization on a set of attributes known to be important for trust, such as
competence and integrity. Participants will also rate how often they communicate about the other members. Participants will receive emails alerting them to complete the evaluations at four points over the course of the year. From the evaluations, I will be able to capture both the favorability (average rating) of each member’s reputation as well as how broad his or her reputation is (how many other members have formed an impression of the person on the given attribute). I will then be able to test my hypothesized patterns of reputation by comparing reputation favorability and breadth for each attribute over time. I will also be able to examine whether communication mediates the relationship. Finally, I will be able to use the information to examine the relationship between reputation and creation of social networks.

**Faculty Advisor**

I am currently working in this area of research for my dissertation, under the advisement of Maurice Schweitzer. With this study, I will be able to address some of the open questions that have stemmed from my dissertation research thus far.

**Budget**

The only funding need I have is in compensating participants for completing the project. Participant compensation: $3200 [$8/period*4 periods*100 participants]

**Other Funding Sources**

I have no other funding sources for this project. Related research has been partially funded by Maurice Schweitzer’s grant and a law firm at which I am gathering data.
Project Title: Significant Interactions: The relationship between conflict types and decision quality.

Applicant: Noah Eisenkraft
(617) 686-4891
nce@wharton.upenn.edu
Department of Management
3620 Locust Walk
Room 3116 SH-DH

Faculty advisor: Associate Professor of Management Sigal Barsade

Project Summary:

Academics have traditionally studied decision processes by examining the biases, abilities, and strategies of individuals. Decisions, however, are not only made by individuals acting alone; people often choose or are compelled to make decisions when they are part of a dyad or group. As they attempt to make decisions, the members of a decision-making unit may disagree and fall into conflict. The research I propose in this application investigates the relationship between conflict and decision-making quality. More specifically, I hope to conduct two studies that replicate, generalize, and extend the results of a three-month study I recently conducted on the effect of conflict on decision-making in entrepreneurial teams.

A major theme of study in the management literature looks at how two different types of conflict—task and relationship conflict—differentially affect performance variables such as decision quality. Jehn (1995) found that relationship conflict would have a negative effect on performance while task conflict would have a positive effect and produced data supporting this assertion. Subsequent research has been far more equivocal. Although the negative relationship between relationship conflict and performance has been supported in a variety of settings, the relationship between task conflict and performance is murky at best; some studies report a positive relationship, others report a negative relationship, and some find no statistical evidence for a positive or negative relationship between task conflict and performance. More recently, researchers interested in studying the relationship between conflict and performance have devoted their energies to studying mediating variables—e.g. trust (Simons & Peterson, 2000)—and new conflict types—e.g. process conflict (Jehn, 1997). These efforts have produced some promising results, but a parsimonious understanding of how conflicts affects decision-making remains elusive.
The data I collected during my aforementioned study suggests that the effects of conflict can be understood without looking for new conflict types or mediators and moderators. Unlike previous research that assessed the impact of conflict by administering post-hoc surveys that ask about conflict over a long period of time, I surveyed the entrepreneurs about their conflicts and decisions every three hours while the teams were at work for 10 weeks. Using a three-hour window allowed me to see whether individual conflict episodes are best described as either relationship conflict or task conflict. I found they could not be; elements of relationship conflict and task conflict were often perceived to exist during the same interaction. If the so-called types of conflict can occur at the same time, incorporating the interaction between task and relationship conflict in the statistical analysis of the data may reveal why previous surveys have reported such a wide range of results.

Incorporating the interaction between task and relationship conflict into my statistical analysis produced results that may provide insight into conflict’s effects for both academics and practitioners. When my data from this summer were analyzed without the interaction effects, the results indicated a strong negative effect for relationship conflict and an uncertain effect of task conflict. When the analyses included the interaction between relationship and task conflict, I found that task conflict in the absence of relationship conflict had a small but significant positive effect on perceived decision quality. When task conflict occurred at the same time as relationship conflict, decision quality dropped precipitously. Somewhat surprisingly, the analyses did not report that relationship conflict alone had an effect on performance.

The results described above suggest that academics may want to rethink their conceptualization of conflict and that practitioners may want to adopt group decision-making strategies that take the nature of the ongoing conflict into account. However, one study cannot produce enough evidence to confidently state these conclusions. For that reason, I will conduct two studies that overcome some of the limitations of the earlier study:

**Study 1:** My study on entrepreneurial teams measured conflict and decisions as they occurred in a natural setting. When measuring naturally occurring phenomena, it is impossible to know whether the relationship between the independent variable and the dependent variable is causal. Study 1 tries to correct for this problem by using an experimental design that sacrifices realism for a higher degree of control.

Study 1 is a survey study that uses an experimental design. Participants in the study will be given a written description of an interaction a regional purchasing manager for a direct sales computer company experiences before making a decision. The description a participant reads will depend on the experimental condition the participant is assigned to. The conditions reflect a 2 (task conflict, no task conflict) x 2 (relationship conflict, no relationship conflict) factorial design. After reading the description of the interaction, the participants will complete a questionnaire that evaluates the task and relationship conflict experienced by the purchasing manager and the predicted perceived quality of the decision.

**Study 2:** My study on entrepreneurial teams measured how the conflict experienced by entrepreneurs at different times affected the perceived quality of the
decisions the entrepreneurial team made. Although there are benefits to using a within-subjects repeated-measure design on an unusual population, this type of study design does not foment confidence in the generalizability of the study’s results. Study 2 measures conflict and decision-making quality using a large sample of dyads/groups to test the generalizability of my previous study’s results.

Like my previous study, Study 2 attempts to measure the effect of conflict immediately after an interaction. To do this, I will invite undergraduates to participate in a negotiation exercise similar to the ones commonly used in negotiation courses. After the interaction, subjects will complete a survey that measures the quality of decisions made during the negotiation and the conflict experienced during the interaction. Study 2 will measure how task and relationship conflict affects both the perceived and objective quality of the decisions made during the negotiation.

Budget:

$500 for a raffle prize to be auctioned off to one of the participants in Study 1.

$1000 to compensate 100 subjects ($10 each) for their participation in Study 2.
Application for Dr Ackoff Doctoral Student Awards for Research on
Human Decision Processes

Doctoral Dissertation in Systems Engineering
University of Pennsylvania

Modeling the Behavior of Human Agents in Virtual Worlds:
A Knowledge Engineering based Systems Methodology for carrying out Integration of Models from the Social Sciences

Prepared for: Dr Robert J. Meyer and Dr Howard Kunreuther
Co-Directors & Professors Risk & Decision Process Center
meyerr@wharton.upenn.edu kunreuth@wharton.upenn.edu
Wharton School University of Pennsylvania

Student: Gnana K Bharathy
Doctoral Candidate in Systems Eng & Ackoff Center for Advancement of Systems Approaches (ACASA)
bharathy@seas.upenn.edu
E-Mail: ESE Dept., 248 Towne Building University of Pennsylvania
Address:

Advisor: Dr. Barry Silverman
Professor
barryg@seas.upenn.edu
E-mail:

Committee: Drs. R. Eidelson, J. Pourdehnad, B. Silverman, T. Smith, L. Starr and I. Zandi
23 September 2005
Date: Final
Status: 1
Version:
1 Descriptive Summary

1.1 Dissertation in General

Although all nuances and complexities of human behavior and the complex interactions in the social systems can not be adequately captured to predict human behavior, the virtual worlds constructed out of realistic human behavior models can assist in exploring the space of possibilities and provide immersive training to decision makers. According to researchers (2003a,b; Pew et.al. 1998, Ritter et.al. 2003, Silverman et.al., 2006a), a common challenge running throughout these applications is to increase the realism of the synthetic agents’ behavior and coping abilities. The demand for the human behavior models is coming from such varied domains as military simulations of immersive training, exploring alternative hypotheses in intelligence analysis, developing friendly and engaging product interfaces as well as domain knowledge acquisition, to name a few. If we were to have realistic and reliable models of virtual worlds, one must find ways to integrate scientific know-how across many disciplines, and to create integrated models from hitherto fragmented frameworks and models in social sciences, particularly psychology and cultural factors as well as physiological constraints to decision making.

In order to demonstrate integrated models of human behavior, I employ an agent framework developed by Silverman et.al. (2001, 2003), which uses bounded rationality for its decision-making, and has provision to accommodate a value hierarchy, represented in the form of a tree or a network. This value hierarchy represents goals (short term needs), standards (e.g. thou shall not do this type constraints) as well as preferences (long term needs) (See Figure 3, in the Appendix). This agent framework would be rich enough to represent cultural values as well as personality traits. The utilities estimated by multiplying up the tree is subjective, expected utilities (SEU- see Figure 2, in the Appendix), as these parameters are subjective estimates that have to be frequently estimated from expert inputs and empirical materials. As per ecological psychology of Gibson, the actions available to the agent are situated on, and are afforded by, the objects (which might include inanimate objects as well as agents) perceived by the agent. Each object might carry multiple perceptions that are visible to an agent under different circumstances described by different rules. Essentially, this is a very smart knowledge based agent framework for modeling human behavior. Silverman et.al. have developed this framework into a software named PMF Serv, which reduces the programming effort. When carrying out human behavior modeling (HBM), it is our experience that one single model in social science/ psychology, cannot provide all the necessary behavior. One has to integrate different models and frameworks from social sciences, physiology, biology etc. Being an open platform, PMF Serv can accommodate different models from the social sciences, physiology etc, and enable integration of different models to create virtual worlds. A brief glimpse of this mathematical model of decision making, has been provided in the Appendix. The software framework mentioned above, like most other human behavior modeling frameworks, relies on external input for incorporating knowledge into the synthetic agents about their cognition, personality, culture, physiology as well as information about the world around these synthetic agents. Typically, the knowledge takes the forms of graphical models constructed (structure), as well as the parameter values ascribed to a large number of model parameters (including weights on the graphical models). Therefore, human behavior modeling (HBM), using such knowledge-based systems software, revolves around the lifecycle of the knowledge models.

Yet, there is no systematic process found in the literature or in practice for constructing, validating and refining models of human behavior, and hence, carrying out human
behavior modeling (HBM) in a defensible manner. The implications of this are diagrammatically described in Figure 1 (Appendix).

The objective of my dissertation is to develop such an integrative modeling process - one that helps developing consolidated models of decision making agents. Specifically, I am interested in structuring the value trees and estimating the parameters of the agents that I model, so that these agents make decisions and exhibit behaviors that are realistic.

Since the available models and frameworks in social science literature are either qualitative, or the available numerical data are limited, empirical materials, the emphasize of the integrative modeling process is on how to: systematically transform empirical evidence, tacit knowledge and expert knowledge into data for modeling; reduce, if not minimize, the human errors and cognitive biases (for example, for confirming evidence); ensure that the uncertainties in the input parameters are addressed; and verify and validate the model as a whole, and the knowledge base in particular.

For the lack of a better term, the process has been conveniently referred to as Knowledge Engineering (KE) process due to extensive involvement of KE techniques and construction of the knowledge models. The research methodology for developing the KE process has been summarized in Figure 4 in the Appendix. The output of the research methodology is the validated KE Process (along with tools and templates). A diagrammatic representation of the KE Process is given in Figure 5 in the Appendix.

In my on-going dissertation, which has been awarded this year’s INCOSE-Stevens award for promising dissertation in Systems Engineering and Integration by International Council on Systems Engineering, I have already developed a systems methodology for integrating existing, but fragmented, frameworks in social science and modeling human behavior through knowledge engineering based process. I have employed the same to create several models of leaders, followers, crowd members, rebels, agitators in conflict situations. I have been able to validate my integrative and modeling KE process under naturalistic conditions by testing, verifying and validating these models. In the future, I hope to expand the work by modeling effects of marketing on consumers, interdependent risk management situations (where some of my previous experience is) and organizational behavior.

While decision-making under uncertainty is one of the threads that run throughout my dissertation, there are two aspects of this knowledge engineering process that are highly relevant to decision making under uncertainty.

Firstly, the overall outcome of this modeling is interested in exploring emergent macro-behavior due to micro-decisions of bounded-rational agents endowed with personality, culture, emotions and incomplete information that lead to limits of rationality. Using the virtual worlds hereby constructed, we explore the landscape of decision space through sensitivity analysis and generate counterfactuals (alternative histories).

Secondly, the aspect that I would like to emphasize is a section in the modeling process, where I have developed several tools and techniques that control the errors and biases of the modelers. I am interested in carrying out a controlled experiment to evaluate the performance of the key aspects of my process.

Among them are tools for differential diagnosis (differential diagnoser) and pairwise comparison, which help elicit parameters in the graphical models through systematic, defensible and transparent process. These tools accompany a mathematical framework, and contain provisions for estimating uncertainties in the expert inputs and empirical evidence. A framework for estimation of uncertainties in using these tools has been given in the Appendix.
When constructing models of behavior from evidence (be it from empirical evidence from literature or expert input), the modeler employs the cognitive process of determining the motives of someone's behavior. This is subject to several biases; the foremost among them are confirmation bias, which might also subsume other biases such as availability bias and attribution bias. Simultaneous evaluation of multiple, competing hypotheses is very difficult to do and is against the cognitive bias of the human mind. Without an instrument for countering this fundamental bias, modelers attempting to build Value Tree Models could be easily misled. The tendency to build models by confirming a plausible but favorite hypothesis will have enormous implications when collating evidences provided by experts and empirical materials and building Value Tree Models. Instead, human mind works though satisfying strategy. The process of selecting the favorite hypothesis is highly influenced by one's own conditioning, and there is the tendency is to see what one is looking for, and to overlook alternatives. Assessing evidence and attributing behavioral traits should ensure that: no external cause explains the same behavior; other competing traits do not explain the same behavior; and confirmation bias is eliminated by disconfirming hypothesis.

In order to minimize the risk of not considering alternatives and considering non-diagnostic evidence, I have provided a tabular design (see Appendix for details, Figure 6 and mathematics), to carry out Differential Diagnosis, for generating and weighting alternative hypothesis, explanations or conclusions. Accordingly, the process forces one to look at competing hypotheses, and disconfirm from an alternative list of hypothesis rather than confirming one hypothesis using available evidence. Once again, testing the usefulness and effectiveness of this differential diagnosis tool will be part of this dissertation. Later, we will look at how this tabular structure can be exploited quantitatively.

The hypotheses in this case are the nodes in the Value Trees (goals, standards and preferences of the characters being assessed). On the left, the framework includes the Organizer containing key evidence, thematically coded and attributed with reliability and relevance. However, the tool for Differential Diagnosis is centered on the hypotheses and evidence. Essentially, the hypotheses are pitted against the evidence through this matrix. If reliable evidence rejects a hypothesis, then the likelihood of that hypothesis is diminished significantly. I advocate including reliability and relevance for each evidence. Relevance (from the Organizer) identify which items are most helpful in judging the relative likelihood of the hypotheses, and help control the time spent on what would seem like irrelevant evidence. Using the Bayesian framework, I have developed a quantitative technique for the differential diagnosis, which attribute higher (an order of magnitude or more) weight to disconfirming evidence. I also established that otherwise rare events weigh higher when they do occur as evidence.

1.2 Proposed Experiment

I have employed the decision aid tools centered around differential diagnosis in our case studies, where several models of leaders, followers and members of the crowd, rebels, such as intifada have been successfully created and tested. I have also carried out a pilot study to investigate the cognitive tasks that people carry out while constructing the models. As part of my research design (see Appendix), I would like to conclude my dissertation by carrying out controlled experiments to evaluate the performance of the differential diagnosis and pairwise comparison process in profiling characters from empirical evidence. The proposal here request funding to support such an experiment.

The objective of the comparative evaluation is to investigate whether the proposed process performs better than the alternatives. The null hypothesis is that the proposed differential diagnostic process (experimental group) does not perform significantly better
than other processes. As a pre-test, the participants will be queried, if they had any background or experience in profiling characters as well as about their general academic performance.

The potential treatments considered are:

- No Treatment: No exposure to formal differential diagnostic and pairwise comparison process or the tools, and participants are left to their own devices, and
- Experimental Treatment: Exposure to actual process containing details of differential diagnostic and pairwise comparison process.

<table>
<thead>
<tr>
<th>Experimental</th>
<th>Pre-Test</th>
<th>X (+CTA)</th>
<th>Post Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Pre-Test</td>
<td>. (+CTA)</td>
<td>Post Test</td>
</tr>
</tbody>
</table>

I have chosen not to include a placebo as control - a treatment arm of rudimentary process containing the same overall steps, but these steps do not have any specific techniques. Typically, in this group, participants are left to their own devices to come up with specifics. It might be argued that introducing no-formal process to a group may be an unfair comparison against a structured process, and it appear to make not much sense to do so.

Small collection of evidence relating to small sections of the value structure of well profiled agents (a leader in business or politics with multinational engagement that we have studied and about whose segment of the value structure, a profile is available) will be handed to the group. The group need not know who the character is, except his/ her actions and contextual evidence will be revealed. All the required information for the creation of the characters will be provided to the groups, and the groups would not know whether the other groups existed (sessions will be conducted separately). The evidence will be organized thematically, but it will not be explicitly marked as such. Subjects for external evaluators would come from the pool of undergraduate/ graduate students at the University of Pennsylvania, recruited through advertisement on campus.

The evaluators (as well as the participants) are blinded to the treatment process, except when evaluating the audit trail, which is carried out after completing other evaluations. Unfortunately, related samples cannot be used, as one hypothesizes that order effects (carry over/ maturation) are expected to matter. Therefore, I propose to match the participants, based on their experience in projects and performance in their coursework, into three classes of achievers, medium performers and non-motivated. Within each group, I will administer the different treatments in equal numbers. The matched participants are then assigned to experimental conditions randomly. In addition to constructing the value structure of the characters, the participants will answer questionnaire to evaluate their cognitive task (cognitive task analysis or CTA).

The sections of the models that participants create will be evaluated for interparticipant-consistency, representational accuracy, audit trail as well as computer implementation in model. The cognitive task analysis through a simple survey instrument will be used to shed light on any exogenous issue interfering in the experiment and what participants themselves have been seeing the process as.
2 Budget

2.1 Description of other sources of research funding

The primary source of support for the researcher support is Beck Fellowship, which covers the tuition and stipend till the end of Fall 05. In addition, the part of the research that is aligned to the projects at the Ackoff Center is supported through the respective project.

If the additional funds requested here in this proposal were made available, I would be able to carry out an experiment to extend the dissertation to compare the performance of my knowledge engineering methodology with ad hoc method that people would use to model these characters.

2.2 Requested Budget

The support requested through Ackoff Award is the cost not directly related to the researcher’s labor as well as materials. Particularly, the researcher would be grateful, if the time of the participants and preferably the evaluators (or part there of) could be provided ($6,400).

2.3 Detailed Budget

Costing for Controlled Evaluation of Differential Diagnosis

<table>
<thead>
<tr>
<th>For estimations:</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher $/hr</td>
<td>$ 20.00</td>
</tr>
<tr>
<td>$/day</td>
<td>$ 160.00</td>
</tr>
<tr>
<td>Participant $/hr</td>
<td>$ 12.00</td>
</tr>
<tr>
<td>$/day</td>
<td>$ 96.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Days</th>
<th># Personnel</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1   Initiation &amp; Authorizations (IRB)</td>
<td>6.0</td>
<td>1.2</td>
<td>$ 1,152.00</td>
</tr>
<tr>
<td>2   Pilot Study (Completed)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development of the Instrument</td>
<td>2.0</td>
<td>1.0</td>
<td>$ 320.00</td>
</tr>
<tr>
<td>Selecting &amp; Recruitment of Participants</td>
<td>0.5</td>
<td>1.0</td>
<td>$ 80.00</td>
</tr>
<tr>
<td>Pilot Trial of the Instrument</td>
<td>0.5</td>
<td>1.0</td>
<td>$ 80.00</td>
</tr>
<tr>
<td>Analysis of Beta Results &amp; Learning</td>
<td>0.5</td>
<td>1.0</td>
<td>$ 80.00</td>
</tr>
<tr>
<td>3   Design of Experiments</td>
<td>2.0</td>
<td>1.0</td>
<td>$ 320.00</td>
</tr>
<tr>
<td>4   Experimental Study &amp; Data Collection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tuning of the Instrument</td>
<td>0.5</td>
<td>1.0</td>
<td>$ 80.00</td>
</tr>
<tr>
<td>Trials &amp; Data Collection (Researcher/ Evaluators)</td>
<td>1.0</td>
<td>4.0</td>
<td>$ 640.00</td>
</tr>
<tr>
<td>Participant Time</td>
<td>1.0</td>
<td>60.0</td>
<td>$ 5,760.00</td>
</tr>
<tr>
<td>Independent Observation &amp; Review</td>
<td>1.0</td>
<td>1.0</td>
<td>$ 160.00</td>
</tr>
<tr>
<td>5   Data Analysis</td>
<td>5.0</td>
<td>1.0</td>
<td>$ 800.00</td>
</tr>
<tr>
<td>6   Reporting &amp; Writing Paper(s)</td>
<td>5.0</td>
<td>1.0</td>
<td>$ 800.00</td>
</tr>
<tr>
<td>7   Presentations</td>
<td>5.0</td>
<td>1.0</td>
<td>$ 800.00</td>
</tr>
<tr>
<td>8   Closure</td>
<td>0.3</td>
<td>1.0</td>
<td>$ 48.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Days</th>
<th># Personnel</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>9   Materials, Supplies &amp; Communication</td>
<td></td>
<td></td>
<td>$ 500.00</td>
</tr>
</tbody>
</table>

$ 15,408.00
Total Cost

$15,908.00
3 CONSEQUENCES OF AD HOC FRAMEWORK INTEGRATION & MODELING

4 DECISION FRAMEWORK FOR AGENTS

4.1 A Model of Decision Making

This section introduces the cognitively detailed decision model through the following state model. Let us assume that one could enumerate the decision paths at time $t$ and $t+1$ as below, where a course of action $COA_{ij}$ take state $S(t,i)$ to State $S(t+1, j)$ with a probability $P_{ij}$. Suppose, a utility $(uij)$ based on possible consequences were to be available for each $COA_{ij}$, expected utility of taking an action might be calculated and taken into account while making the decision. The agent selects the best course of action based on the available information and by maximizing the expected utility, $E_i$ (calculated as sum across outcomes of utility times probability).

---

**Figure 1: Consequences of Not Having a Process for Modeling**

**as-it-should-be**
- valid models, learning from and leveraging past experience,
- using appropriate sources of information,
- systematic, lifecycle process, models, justifiable, transparency,
- techniques to minimize modeling errors and biases,
- quantification of uncertainties, verification, validation,
- sensitivity analysis,
- opportunity for learning & mgmt of KBs

**as-it-will-be**
- ad-hoc models, not validated,
- modeling errors, biases,
- No learning from experience

**as-it-is**
- Composable, knowledge-based software available
- No adequate methodology

---
Figure 2 – State Model of COAij available at State S(t, i)
In order to apply the State Model, we make following simplifications, by assuming independence between the state model, reducing the multitude of outcomes that are possible as a consequence of a given action to discrete set of defined outcomes, and limiting the number of steps an agent looks ahead (to say one). Since agents interact with each other, as agents are making similar decisions based on beliefs that options lead to a given next state, they are dependent on each other to perform up to par for that next state to be reached. In reality, individual state models are not completely independent, as at any given time each COA consists of a work packet to be performed by one or more agents in the presence of opponent actions and other internal and external moderators. However, in local regions (k), the interactions between these models may be treated as perturbations, allowing the problem of determining the overall Z(t) to be decomposed into simpler problems of determining individual z_k(t) and combining them to obtain Z(t).

Now let us look at how each option or COAij might be represented in PMF Serv architecture. The following is a model of Silverman’s (2002a, b) PMF Serv Agents. For every agent in PMF Serv, the following are specified: value trees of goals, standards & preferences (GSP), weights for the arcs on the trees, and inputs regarding stylized physiology (tank models and parameters). For the Micro world in PMF Serv, stereotypical views of agents & objects along with actions and potential consequences are also specified.

Consider the following schematic diagrams referring to representing cognition and personality in agents as a value hierarchy. Such as value hierarchy (or other graphical models with some modification) may be used to model the decision making by the agent.

Figure 3: Actions Triggering Success/ Failure in Value Hierarchy
There exists a value hierarchy in every agent. The nodes of the value tree are goals to be pursued, standards of behavior to be adhered to, and preferences (likes and long term objectives). These nodes are weighted based on their importance. Before a given action is taken by an agent, the agent evaluates the available alternatives by assessing the success or failure of the value nodes, when a given action is taken. Such events as physiology and stress serve as constraints on the ability make decisions or act on their decision. Similarly, the perceptual types are stereotypical views of the world around the agent, which afford agents actions. For example, viewing an object as a waste paper basket gives agent the choice of throwing waste paper inside the object.

An action, when attempted, could succeed (or fail) with a certain likelihood (P). Both success as well as failure of the attempted action separately results in the failure or success of various leaf nodes (depending on the pre-determined agreement or discord between the action outcome and the leaf-node values). In order to record the failure and
success of the nodes, on each node of the Value Tree Models are two “reservoirs” that reflect the current activation of success and failure of this node, respectively.

The various situations, which exist in the current state of the world and consist of valence and intensity, triggers success or failure up the Value Tree Model, thereby propagating utility estimates. Undertaking an action (as well as contemplation) triggers success of failure of the appropriate tree nodes (S & F), which in turn triggers emotions with respect to an action. Activation of a node leads to the weights (Wis) being multiplied together up the branch of the tree from leaf node to root (as Wijl(c) = PWij) to compute the concern. The expectation of the concern over the entire set of trees, taking into account the likelihood of the action succeeding and failing, is simplified as the Intensity (ξ).

The expected emotional consequence of the action, including the subjective estimate of the likelihood of success of the action, form subjective utility. Under normal operating conditions, the agent selects the COAj by maximizing this subjective utility, based on the strength of their beliefs (Expected Values) that the next state is on the path to maximum utility.

In general, any of a number of k diverse activations could arise with intensity, ξk, and that this intensity would be somehow correlated to importance of one’s values or node set and whether those concerns succeed or fail for the state in question. This can be expressed as:

\[
ξ_k(b \in B) = \sum_{j \in J_k} \sum_{v \in V} \left[ Wij(v \in V) \ast φ(r_j) \ast ζ(v) \ast ψ \right]
\]

Where,

| \(ξ_k \rightarrow ξ_k(b \in B)\) | Intensity of activation, k, due to the bth state of the world. |
| Jk | The set of all agents and objects relevant to k. J1 is the set consisting only of the self, and J2 is the set consisting of everyone but the self, and J is the union of J1 and J2. |
| W(v \in V) | Weighted importance of concern V to the agent. |
| V | The set of goals, standards, and preferences held by the agent. |
| Φ(rj) | A function that captures the strength of positive and negative relationships one has with agent or object j that are effected or spared in state b. This is estimated as relationship between agents. |
| ζ(v) | Degree of activation for a goal, standard, or preference due a given event. A degree of activation is defined on the scale of 0 to 1 (for both success and failure of the leaf node) and relates the type of action in the world and the strength of their relationship to events. |
| ψ | A function that captures temporal factors of the state and how to discount (decay) and merge one’s GSP activations from the past (history vector), in the present, and for the future. |

It is important to note that the weights adhere to principles of probability; e.g., all child node insights add to unity beneath a given parent, activations and weights are multiplied up a branch, and no child has multiple parents (independence). Although we use fixed weights on the Value Tree Models, the reservoirs serve to render them dynamic and adaptive to the agent's current needs. Thus, when a given success reservoir is filled, that tends to nullify the importance of the weight on that node (or amplify it if the failure reservoir is filled). In this fashion, one can think of a form of spreading activation (and deactivation) across the value structure as the scenario proceeds.
As one can see, it is the activation, interaction, and decay of parameters on trees and reservoirs from the first-person perspective of the agent, not a rule or schema coded by a knowledge engineer in the third-person, that allows the agent to try and maximize utility to decide its next course of action. The result is an agent attuned to the environment that adapts and responds to different situations subject to its perception, value system, relationships, and needs.

5 HOW MY RESEARCH WILL CONTRIBUTE TO HUMAN DECISION MAKING

5.1 Significance of the Dissertation

The significance of the research is also tied to the significance of being able to use cognitively detailed (thick) agents in modeling. The knowledge engineering process can help construct knowledge-based agents with cognition and personality in a justifiable manner, furthering the use of HBMs, in decision support, immersive training and knowledge acquisition.

5.2 Salient Contributions

The salient contributions of the dissertation include:

- taking a systems perspective, and identifying and questioning the purpose, context, and boundary for the work;
- demonstrating the integration of diverse models and frameworks from social sciences and physiology (Contribution: Integration);
- integration of a collection of existing techniques tailored for eliciting input data systematically from empirical evidence in literature, experts and user groups, with particular contribution in the form of integration of different approaches (Contribution: Integration);
- developing tools and techniques for organizing evidence through a thematic scheme, and eliciting weights through a pairwise comparison process (Brannon, 2002; Xu, 2003), and implementing a differential diagnosis to disconfirm evidence (Dunbar, 1993; Popper, 1959; Gorman & Gorman, 1984) (Contributions: Techniques themselves for building the knowledge base);
- integration of a realistic validation process based on triangulation of number of appropriate techniques (Contribution: Integration);
- integration of existing techniques for addressing the uncertainty issues pertaining to weight and other parameter determination (Contribution: Integration); and
- developing a collection of paper-based tools to support various steps in the process (Contribution: Tool development).

6 RESEARCH DESIGN

The key steps for the development of the KE process are outlined in the research design (see Figure 2). At this juncture, for the sake of clarity, one must also distinguish between
the research methodology that is being proposed for development of the KE process, and the KE process itself.

Figure 4: Dissertation Research Design

7 **Knowledge Engineering based Integrative & Modeling Process**

A brief description of the KE based integrating and modeling process is given below (Figure 3). The knowledge engineering process would help identify and integrate different models, construct the value tree of the agents with goals, standards and preferences to reflect the cognition, personality and socio-cultural conditioning of the agent, systematically transform information from diverse sources (particularly the empirical evidence but also expert input) into data for modeling, assign weights to the value structure nodes, and construct the objects in the world. The model thus constructed will be verified and validated.
Conceptualization & Planning
- Understanding the Scope, Containing System
- Selecting the Level of Model Complexity & Appropriate Level of Detail in the Model
- Understanding the Models to work/Integrate with
- Scoping the Corresponding Design Issues
- Developing Appropriate Integration Interface

Requirements Analysis & Drill Down
- Characterizing the Domain and the HBM Modeling Problem
- Selecting Appropriate Social Science Frameworks, if available
- Identifying and Assessing the Key Variables (5 Ps)
- Developing the Baseline Scenario & Deciding on the mix of specific individuals & archetypes

Acquiring & Organizing Information
- Collecting Evidence on Population & Societal Studies
  - Experts, Users, (or from the field studies through experts)

Developing Cognitive, Personality & Culture Structure (GSP Trees)
- Identifying the Dimensions Required to be Modeled
- Constructing Models by Laddering
- Dimensionality Reduction
- Designing the Archetypical Personality to be Represented

Assessing Weights (Parameters)
- Confirming the Required Traits using Societal & Cultural Evidence
- Disconfirming Unwanted Traits against Textbook Archetype and Societal Evidence

Developing the World of Objects & P-Types using DI Model
- Spiral Refinement, & Reusing
- Monitoring

Uncertainty

Creating the Virtual Social Space & Integrating as Necessary
- Creating Scenarios & Objects
- Introducing Specific Characters
- Creating a Distribution of Characters from Crowds
- Carrying out Software Integration (as Necessary) & the Integration Testing

Model Verification & Validation

Sensitivity Analysis & Exploration of the Decision Space

Management of Knowledge Bases

Figure 5: KE based Process
8  Differential Diagnosis & Pair wise Comparison Process

The following are some of the key steps in my process, where decision making under uncertainty plays a role. The specific interest is in differential diagnosis and pairwise comparison process around which the controlled experiment is proposed.

8.1 Organizing Empirical Materials

Organizing the evidence is an important step that requires a systematic process, because the organized information from otherwise diverse or amalgamated sources is critical to the success of the remaining knowledge engineering activities. In order to help organize the data, I have employed a modified content analysis process (Krippendorf, 1980). In order to do this, the relevant themes are obtained from the information available on the domain, including the high-level goals, standards and preferences if available. The body of materials would be split up into records with each representing one and only one theme. Then, these are assigned theme codes, relevance and reliability, and sorted according to the themes.

8.2 Differential Diagnosis

Humans tend to confirm a hypothesis and commonly accord more weight to evidence supporting a hypothesis than to evidence that weakens it. In addition, a number of similar cognitive limitations in humans have been documented by the researchers (Tversky and Kahneman, 1974; Bruner, Goodnow and Austin, 1956). At the same time, it was proposed that conclusive verification of hypotheses is not possible; only conclusive falsification is possible, and thus participants should seek to disprove favored explanations of various phenomena (Popper, 1959).

When constructing synthetic characters, there will neither be sufficient data nor time to carry out an exhaustive analysis. Although qualitative and semi-quantitative analysis such as one carried out in this study do not lend themselves to statistical procedures of scientific methodology for users to test their hypotheses, they can and should adopt the conceptual strategy of seeking to refute rather than confirm hypotheses. This is the basis of the differential diagnosis devised and employed in this thesis, where evidences are used to disconfirm hypothesis than confirm them.
8.2.1 A Framework for Differential Diagnosis

Our goals here are to:

- develop a quantitative technique for the differential diagnosis and disconfirming evidence in the Bayesian framework and reliability,
- generalize the formulae for multiple evidence and hypothesis, and
- relate the confidence index, a simple expression for uncertainty, to that derivable through formal method.

We have seen that the natural human inclination is to find support in the evidence. The extreme case is that, in the empty cells, most evidence can be construed as supportive. As one can see, evidence supporting a hypothesis is not very powerful, as the same evidence can support multiple hypotheses (may not have much diagnostic value). It is evidence that can be used to reject a hypothesis that provides the real power to this technique. Nevertheless, evidence confirming a subgroup of (but not all) competing hypothesis might also provide weak capability in distinguishing these groups of evidence from the rest. Therefore, disconfirmation is much more powerful technique, confirmation also provide some weak diagnosis. We also find that evidences available may or may not be reliable.

These have two implications:

1. Instead of categorically rejecting a hypothesis when the evidence is contrary to the hypothesis, it is possible to take into account the strengths and reliabilities of the evidences and probe further.
2. The evidence supporting a hypothesis is not as powerful as the evidence that can be used to reject a hypothesis, and must be weighed lower than the latter.

Using this information, it is possible to derive semi-quantitative estimates from a differential diagnostic a table. Suppose, instead of rejecting or accepting a hypothesis, let us consider a set of (n-number of) evidences relevant to a set of competing hypotheses. Let us assume that the evidence (Ei) with a reliability Ri, rejects (or supports) a hypothesis (Hj) with a strength (Cij), where Cij ∈ (-1, +1). Cij value of +1 implies full support, while –1 implies complete rejection, as assessed by the expert/ knowledge engineer. Now, let us place this “arbitrary” strength of confirmation (Cij), and hence differential diagnosis, in the Bayesian context. We know that Bayes’ Theorem can be expressed as:

\[
P(H|E, c) = \frac{P(H|c) \cdot P(E|H, c)}{P(E|c)}
\]

where we can update our belief in hypothesis H given the additional evidence E and the background context c. The left-hand term, P(H|E, c) is known as the “posterior probability,” or the probability of H after considering the effect of E on c. The term P(H|c) is called the “prior probability” of H given c alone. The term P(E|H, c) is called the “likelihood” and gives the probability of the evidence assuming the hypothesis H and the background information c is true. Finally, the last term P(E|c) is independent of H and can be regarded as a normalizing or scaling factor.

It is important to note that all of these probabilities are conditional. They specify the degree of belief in some proposition or propositions based on the assumption that some other propositions are true. As such, the theory has no meaning without prior resolution of the probability of these antecedent propositions.

In the given context c, P(H| c) measures ones degree of belief in a hypothesis when one does not know the evidence E, and the conditional probability P(H|E, c) measures one degree of belief in H when E is known.

E confirms or supports H when P(H|E, c) > P(H| c)  --------[A1.02]

E disconfirms H when P(H|E, c) < P(H| c)  --------[A1.03]

E is neutral with respect to H when P(H|E, c) = P(H| c)  --------[A1.04]

Suppose one were to find out whether Richard is conquering because of his interest in wealth. Not knowing Richard well, no one really knows. So if H3 is the statement "Richard conquered for economic reasons", then P(H3) is taken as around 0.5. Now, Richard selling a territory that he had conquered or plundering a territory gives some support to this hypothesis H3, although it does not prove that H3 is true. The evidence increases your confidence in H, so indeed P(H|E, c) > P(H|c). At the same time, there is another evidence that Richard spends excessively in battles seem to disconfirm the hypothesis. On the other hand, knowing that Richard conquered several territories probably does not make a difference to ones degree of belief in H.

Now, let us define a possible measure of the amount of confirmation (C) as:

\[
C = P(H|E, c) - P(H|c)
\]

The higher the value, the bigger the confirmation. The famous Bayes theorem says:
\[
P(H|E, c) = P(E|H, c)P(H|c)/P(E|c) \quad \text{--------[A1.06]}
\]

Generalizing the equation [13.05] for \(i\)th evidence and \(j\)th hypothesis, the amount of confirmation (\(C_{ij}\)) of hypothesis \(H_j\) by evidence \(E_i\) given the context \(c\), is given by:
\[
C_{ij} = P(H_j|E_i, c) - P(H_j|c) \quad \text{--------[A1.07]}
\]
Using Bayes theorem,
\[
C_{ij} = P(E_i|H_j, c)P(H_j|c)/P(E_i|c) - P(H_j|c) \quad \text{--------[A1.08]}
\]
\[
C_{ij} = P(H_j|c) \{ P(E_i|H_j, c)/P(E_i|c) - 1 \} \quad \text{--------[A1.09]}
\]

Notice that all else being equal, the degree of confirmation increases when \(P(E_i|c)\) decreases. In other words, if the evidence is rather unlikely to happen, this provides a higher amount of confirmation. This accords with the intuition that surprising predictions provide more confirmation than commonplace predictions. So this intuition can actually be justified within the Bayesian framework.

The likelihood of occurrence of an evidence \(P(E_i|c)\) should not be confused with the reliability of the evidence \(R(E_i)\) or simply \(R_i\). \(R(E_i)\) refers to the probability that the evidence \(E_i\) is correctly recorded in the history. The reliability refers to the quality of the information, while likelihood \(P(E_i|c)\) relates to the physical occurrence of the event. The reliability of the evidence should also be factored into account, when the total uncertainty is determined.

As we have seen before, Confidence Index (CI) for a given Hypothesis (\(H_j\)) may be given by:
\[
CI_{\text{Avg}}(H_j) = \frac{1}{n} \times \sum_{i=1}^{n} K \times C_{ij} \times R_i
\]
where \(K = \{ w_1 \text{ when } C_{ij} \geq 0, \text{ and } w_2 \text{ when } C_{ij} < 0 \}. \) Essentially, \(K\) is used to assign a higher weight (say an order of magnitude) to disconfirming evidence (\(w_2 \gg w_1\)).

An expression of this form gives a new meaning to content analysis process, where the contents are treated as passive pieces of information without regard for the uncertainty of the information. This framework also gives a novel tool for disconfirming hypothesis in simple, yet scientifically sound manner. More than the answers themselves, the important part is the systematic and transparent process, which forces an analyst/modeler/knowledge engineer to go through.

The Confidence Index and Confidence Index Ratio can be employed to identify parts of the Value Tree Model where the confidence is low and require further investigation. The investigation may be seeking more information, or determining the sensitivity of the model to available information. Both intra-expert uncertainties in comparison as well as inter-expert differences could be incorporated by considering that the elements of the matrix \([C_{ij}]\) as a random variable \([C_{ijk}]\) where \(k\) is the index for the sample number. The average (wrt to evidence \(i\)) Confidence Index of the hypothesis vector \([CI_{\text{Avg}}(H_{jk})]\) can be estimated using this matrix \([C_{ijk}]\). Both expectation \(E[CI_{\text{Avg}}(H_{jk})]\) and variance \(V[CI(H_{jk})]\) of the Confidence Index can also be estimated from \([CI_{\text{Avg}}(H_{jk})]\). The estimated variance would directly feed into the uncertainty pertaining to hypothesis \(H_j\) while expectation would serve as the estimate of \(CI_{\text{Avg}}(H_j)\). The competing hypothesis that has the highest
confidence clearly wins, if the marginal lead is said to be significant and the variance is small. Otherwise, multiple competing hypotheses might have to be entertained during the course of the sensitivity analysis.

8.2.2 **Consolidating Model(s) from Multiple Experts & Quantifying Variations in the GSP Structure**

In order to elicit knowledge from multiple experts, the knowledge engineer may invite the experts to a meeting and record the resulting outcome of the meeting. Alternatively, the knowledge engineer may consult each expert individually. This process will be handled by communication protocol established in the beginning. However, immaterial of which technique is adopted, a transparent technique for consolidation is necessary. The following section illustrates how this could be carried out.

Inherently, these techniques assumes that each individual experts assessment as an independent sample, the goal is to use the samples to make judgments about what the true parameters of the distribution are. This approach might be particularly useful, if large number of expert opinions is available on the same parameter.

One important aspect that existing techniques fail to mention is that when multiple experts are involved, one must confirm that all experts have the same understanding and the assumptions they make are transparent. These assumptions must be explored in the presence of the expert and the expert is allowed to provide estimates with and without the assumptions involved. A computer interface for exploring the assumptions and revising their estimates without the loss of face (as some experts might believe), might be ideal.

When an expert is allowed to revise his or her standard and when many such expert inputs are combined, the samples might be clustered, and together these characterize the variation within expert as well as between expert. These outputs could be construed as stratified samples with stratification occurring in the samples within expert.

**Consolidation of Numerical Estimates:** When the expert opinion results in the form of numerical values, one of the means of consolidating them could be to treat each value as a random variable and aggregate them through statistical inference.

**Consolidation of Graphical Inputs:** If the expert knowledge can be captured in graphical form, then a different means of consolidating might be necessary. Banks and Carley estimated the true information flow pattern in an organization using the classical statistical inference, employing as data employees’ individual beliefs (expressed as networks) about how the information flows. Massey and Wallace (1991) report techniques that depend upon the use of groups (e.g, focus groups) to elicit and represent knowledge from multiple experts for use in a decision support system. Rush and Wallace (1997) have also demonstrated that influence diagrams from multiple experts may be combined to provide a multiple expert influence diagram (MEID). The same concept could be extended to other type of graphical inputs as well. However, the key difference is that instead of one assuming every input as random, the experts could dialogue with oneself, exposing the assumptions and re-estimating again. Thereby, clusters of understanding from different experts will become available. This method calls for each expert to represent his/her understanding of the decision situation as a directed graph G such as Value Tree Model. Directed edges connect those nodes of the graph. A graph can be represented by an incidence matrix with (i, j)-th element denoted by xij, where xij is the strength of the arc from i to j.

<table>
<thead>
<tr>
<th></th>
<th>Node j</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Z</td>
<td>1</td>
<td>0</td>
<td>x12</td>
<td>x13</td>
<td>0</td>
</tr>
</tbody>
</table>

Graph representation of the incidence matrix:

- **A**\(\rightarrow\)N1
- **B**\(\rightarrow\)N2
- **C**\(\rightarrow\)N3
- **D**\(\rightarrow\)N4
- **E**\(\rightarrow\)N5
Typically, \( x_{ij} \) represents the presence of an arc (Rush and Wallace, 1997). However, it could also represent the weights of the arc, which is the strength of the relationship, where a missing arc could be denoted by zero weight.

Referring to the incidence matrices \( W_1 \) and \( W_2 \) of two Graphs \( G_1 \) and \( G_2 \) respectively and assuming that \( x_{ij} \) indicates the presence or absence of the arc, one could define a symmetric difference matrix, which describes the amount of discrepancies between the graphs \( G_1 \) and \( G_2 \), as below:

\[
d(W_1, W_2) = \text{Tr} [(W_1-W_2)^T(W_1-W_2)] \\
\]

\[
d(W_1, W_2) = 0, \text{ iff } W_1 = W_2 \text{ (reflexive)} \\
\]

\[
d(W_1, W_2) = d(W_2,W_1) \text{ (symmetric)} \\
\]

\[
d(W_1, W_3) + D(W_3,W_2) \geq d(W_1, W_2) \text{ (triangle inequality)} \\
\]

According to Bank and Carley (1994),

\[
P(w) = c(s) \exp (-s \cdot d(w,W)) \text{ for each matrix } w \text{ in } W \\
\]

Where \( s \) is a dispersion parameter and \( c(s) \) is a normalizing constant.

\( W^* \) is initially constructed by adopting a rule of consensus such as: if more than 50% of the GSP Trees have a particular arc, then this arc should be included in the Value Tree Model. \( W^* \) and \( s^* \) (estimates of \( W \) and \( s \) respectively) can be estimated using classical likelihood technique.

\[
L[W, s] = n \cdot \log c(s) - s \cdot \Sigma d(w, W) \text{ for } s \in W \\
\]

\( L[W, s] \) is maximized, when \( \Sigma d(w, W) \text{ for } s \in W \) is maximized.

\[
s^* = -\ln\{(r-1)Z\}/\{1-(r-1)-1Z\} \\
\]

Where \( r = m(m-1) \) given that \( m \) is the number of nodes. \( Z \) is \( \Sigma d(w, W) \), \( \forall w \in W \).

\( s^* \) denotes a measure of disagreement among the experts (or between the estimates of the same expert), and a related measure is the coefficient of dispersion given by \( s^*/m \).

There are few, but significant, implications to these estimations:

**Introspection, Revision and Dialogue**: The degree of disagreement can be used to generate feedback to the experts themselves and their assumptions can be made transparent with subsequent interview. Then the expert can redesign the Value Tree Model in the absence of the assumption. Such an exercise can be brought to the group sessions to discuss the differences. In essence, it can create introspection and dialogue, which could get into the root of the actual problem being studied. In a more superficial treatment, a structure could be adopted by consensus seeking process, or by bootstrapping, or differently weighing expert and lay designs.

**Uncertainty Estimation**: The estimates can also provide estimates of the uncertainty (or confidence) in the Value Tree Model. Rush and Wallace have employed Banks and
Carley’s non-parametric, bootstrap technique for estimation of the confidence interval for influence diagrams.

8.3 Pair wise Comparison Process for Weight Estimation

Incomplete Information: The information available is seldom perfect or complete. However, models have to be constructed with the limited information. In these situations, the weights could be assessed through subjective means, but these models must be used to aid further study and knowledge acquisition. In rare circumstances, the subjective assessment of weights might just suffice.

Consolidating Weights from Different Sources: The circumstances such as these might exist, for example, when an expert is able to provide the weights for parts of the tree, or when different experts provide estimates for different parts of the tree, knowledge engineer must weigh the relative importance of different branches in the same hierarchical level. Experts may estimate the strength of different traits, however, these estimates should be brought together. Through subjective means, pair wise comparison comes to aid to comparatively consolidate disparate weights. This process leaves an audit trail, which can be subject to further investigation and the model further improved when additional information is available.

In both cases, there are no suitable methods for carrying out this in a systematic manner, controlling the error as much as possible. In light of this, I propose a technique (actually a modification of the existing technique) that will reduce the error involved significantly and leave an audit trail for future improvement.

The KE incorporates an AHP (Saaty, 1982) based assessment of weights in GSP arcs. Incorporation of pair-wise comparison caters to the fact that at a given time, human mind can comfortably and reliably compares only two attributes. Pair-wise comparison and AHP based assessment also provide more systematic processes for assessment of weights and leaves an audit trail in the process.

8.3.1 Weight Elicitation through Pairwise Comparison Process – Proposed Technique

The weights of the arcs are semi-quantitatively assessed against every other sibling arcs at the same level, through a pair-wise comparison process. The assessment process itself is subjective and involves pair-wise comparison through AHP scoring scheme. The assessment questions are designed as a structured interview by using the arcs (hence the name of the node below them) to be pitted against each other. The question takes the form, “which of the two attributes are more important to this character”. The knowledge provider, or the expert, can choose along a linear scale.

The process was well illustrated getting a group of students to repeat the procedure. The process was found to be helpful, but further work is being carried out to confirm the effectiveness of this. The method, which is time consuming and rigorous, may be progressively relaxed down the tree. The appropriateness of using this method must be weighed against the likelihood of inconsistency (directly dependent on number of child nodes). Therefore, lower down the tree, one could directly assign weight based on the evidence available. This is supported by the fact that AHP is valuable, only when weights have to be distributed among several competing nodes, and would be commensurate with the depth of analysis needed.

The reader might recollect from the earlier section, the pairwise comparison being an innate trait (Robson, 2001; Cosmides and Tooby, 1994; Samuelson, 2004; Sato, 2004). The scale of 1-9 also reflects the to degree to which one can discriminate the intensity of relationship between elements (Saaty, 1982).
Table 2: Score Description

<table>
<thead>
<tr>
<th>Level</th>
<th>Descriptor</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Equally important or preferred</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Slightly more important or preferred</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Strongly important or preferred</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>Very strongly important or preferred</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>Extremely important or preferred</td>
<td>9</td>
</tr>
</tbody>
</table>

The scores are assigned for each pair of attribute using the following scoring scheme similar to AHP:

However, the total weight given to an attribute in the given node can have a maximum weight of +1 and the minimum weight of 0. Therefore, normalized values of the row averages are used as the weight inputs to the trees.

Consider the stylized case, where weights are being assessed for an agent’s goal tree. The following illustrate a small section in a typical questionnaire, where just two nodes in a given level are being considered at a time. Similar questions are posed for other pairs of attributes. This system for assessment of the weight permits a reduction of many evaluations to a long series of pair-wise comparisons, in which the accumulating results are stored for later calculation while the user can focus serially on distinguishing only two qualities or quantities at any one time.

What would this Character prefer, or find important, between:

<table>
<thead>
<tr>
<th>Love &amp; Belonging</th>
<th>and</th>
<th>Esteem</th>
<th>I am not sure</th>
</tr>
</thead>
</table>

How much more would this Character prefer that? [Or How much more important would this be to this Character? Explain]

<table>
<thead>
<tr>
<th>Equally</th>
<th>Slightly</th>
<th>Strongly</th>
<th>Very Strongly</th>
<th>Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Physiological Needs</th>
<th>Safety &amp; Security</th>
<th>Love &amp; Belonging</th>
<th>Esteem</th>
<th>Self-Actualization</th>
<th>Geo Avg</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physiological Needs</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>4.4596</td>
<td>0.5670</td>
</tr>
<tr>
<td>Safety &amp; Security</td>
<td>1/4</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>1.6227</td>
<td>0.2063</td>
</tr>
<tr>
<td>Love &amp; Belonging</td>
<td>1/7</td>
<td>1/3</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>1.0355</td>
<td>0.1317</td>
</tr>
<tr>
<td>Esteem</td>
<td>1/7</td>
<td>1/3</td>
<td>1/5</td>
<td>1</td>
<td>2</td>
<td>0.4529</td>
<td>0.0576</td>
</tr>
<tr>
<td>Self-Actualization</td>
<td>1/9</td>
<td>1/5</td>
<td>1/5</td>
<td>1/2</td>
<td>1</td>
<td>0.2947</td>
<td>0.0375</td>
</tr>
</tbody>
</table>
Many methods to derive the weight vector from the pairwise comparison matrix have been proposed in the literature, including the mean of the normalised values (an approximate method), the eigenvalue methods (right, left, general), the geometric mean of the rows and the geometric mean of the columns (here we have used geometric mean for simplicity). The following subsection discusses how eigenvalue technique could be exploited to improve consistency. It also provides a mathematical framework/model for incorporating multi-expert input and the uncertainty arising therein. Given that the weights themselves are subjective estimates, the intra-expert uncertainty in comparison decisions and inter-expert differences in estimations could be directly captured as the uncertainty in the weights estimated, using regression approach incorporating random variables. The consolidated uncertainty, in the form of variance and distribution of the weights, could be utilized in the sensitivity analysis.

8.3.2 Ensuring Consistency in the Inputs

Use of AHP based scheme offers an additional advantage that consistency in the assessment can be checked through a procedure outlined by Saaty (1982).

The assessment and consequent assignment of scores in an AHP based scheme can be summarized through a two-dimensional comparison matrix A. Consider such an AHP-matrix of size $m \times m$ with the entries $a_{ij}$, $i, j \in (1, m)$, where $i$ and $j$ are child nodes on a given branch of the Value Tree Model and the weights are designated using this child node identities for convenience. So, $a_{ij}$ is the relative value of attribute $i$ compared to attribute $j$ as perceived by the expert. Many methods to derive the weight vector from the pairwise comparison matrix have been proposed in the literature, including the mean of the normalised values (an approximate method), the eigenvalue methods (right, left, general), the geometric mean of the rows and the geometric mean of the columns. These can be summarized as $\psi(A) = w$ where $\psi(.)$ is a derivation method and $w$ is the weight vector.

The standard method to solve the values for the weights from an AHP-matrix is to take the eigenvector corresponding to the principal eigenvalue of the matrix, and then to standardize the sum of the components equal to one (Saaty, 1982). Geometric and arithmetic means are attempts to approximate this $\lambda_{\text{max}}$. 

Figure 9: Section of the Value Tree

<table>
<thead>
<tr>
<th>Goals Tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physiological Needs</td>
</tr>
<tr>
<td>Safety &amp; Security</td>
</tr>
<tr>
<td>Love &amp; Belonging</td>
</tr>
<tr>
<td>Esteem</td>
</tr>
<tr>
<td>Self-Actualization</td>
</tr>
</tbody>
</table>

Figure 10: Pairwise Comparison Process based on AHP (Saaty, 1982)
\[ \mathbf{A} \cdot \mathbf{w} = \lambda \cdot \mathbf{w} \]  
where \( \mathbf{A} \) is the comparison matrix  
\( \mathbf{w} \) is the weight vector and the eigenvector  
\( \lambda \) is the dimension of the matrix and the eigenvalue

These mathematical solutions also yield estimates of data consistency within the matrix. This matrix is said to be consistent if all their elements \( a_{ij} \) respect the following transitivity and reciprocity rules:

\[ a_{ij} = a_{ik} \cdot a_{kj} \]  
where \( i, j \) and \( k \) are child nodes on a given branch of the Value Tree Model  
\[ a_{ij} = 1 / a_{ji} \]  
\[ a_{ij} = w_i / w_j \quad \forall a_{ij} \]  
where \( w_i \) is the weight of the alternative \( i \)

Consistency may be measured by Consistency Index (ConI) (Saaty, 1982).

\[ \text{ConI} = \frac{\lambda_{\text{max}} - n}{n - 1} \]  
where \( n \) is the number of elements

\[ \text{CR} = \frac{\text{CalcConI}}{\text{RandConI}} \]  
where CR is the consistency ratio (CR \( \leq 0.05 \) is considered desirable)

CalcConI is the calculated ConI  
RandConI is the random ConI for the given matrix, estimated assuming that the assignments were completely random based on the scoring scheme.

We have provided to consolidate the weight inputs from multiple experts as well as to account for variation in the successive estimates from the same individual.

Use of AHP based scheme offers an additional advantage that consistency in the assessment can be checked through a procedure outlined by Saaty (1982). The consistency is important because it gives an opportunity to identify a potential error in the input values. In the absence of error, it may also characterize the lack of certainty in the estimate.

8.3.3 Consolidating & Handling Intra- and Inter- Expert Variations in Weight Assignment

A statistical approach is needed to account for these variations in the input values. Some researchers have applied regression methods to estimate the weights from matrices (Alho et al., 1996; Laininen and Hamalainen, 1999; Crawford and Williams, 1985). The author borrows the same concept to arrive at the estimates.

Let \( v_1, \ldots, v_m \) be the un-normalized values of the attributes which when normalized can yield the actual weights \( w_1, \ldots, w_m \) respectively (\( w_1 + \ldots + w_m = 1 \)). From Equation A1-20, we know that \( a_{ij} = w_i / w_j = v_i / v_j \). Taking logarithm on both the sides, we have for the total of \( m(m-1)/2 \) comparisons made:

\[ \log [a_{ij}] = \log[v_i] - \log[v_j] + X_{ij} \]  
Where \( X_{ij} \) are random variables with expectation \( \mathbb{E}[X_{ij}] = 0 \). This RV \( X_{ij} \) accounts for the potential variation in the estimations and deviations.
from the true value.

For convenience, substituting $\theta_i = \log[v_i] \forall \in (1, m)$, yielding:

$$\log[a_{ij}] = \theta_i - \theta_j + X_{ij}$$

Where $X_{ij}$ are random variables with expectation $E[X_{ij}] = 0$. This RV $X_{ij}$ accounts for the potential variation in the estimations and deviations from the true value.

If one makes an assumption that the random variable $X_{ij}$ is independent and normally distributed with $E[X_{ij}] = 0$ and a common variance $V[X_{ij}] = \sigma^2$, the parameter $\theta_i$ can be solved using Equation [A1.24]. The interesting observation is that $\log[a_{ij}]$ is a random variable with parameters $N(\theta_i - \theta_j, \sigma^2)$. Regression analysis is expected to solve $\log[a_{ij}]$ as a function of $\theta_i$. For example, least squares could be applied after appropriate manipulations such as normalization ($v_m=1$ and $\log[v_m]=\theta_m=0$).

The solution can be computed by using a standard regression program. The estimates of the weights $w_i$ are calculated as $<w_i> = f(<\theta_1>, <\theta_2>, \ldots, <\theta_m>)$ where $<w_i>, <\theta_1>, <\theta_2>, \ldots, <\theta_m>$ are respectively the estimates of $w_i$, $\theta_1$, $\theta_2$, $\ldots$, $\theta_m$ given by the regression analysis. The variances of the estimates $w_i$ can be calculated by using the delta method like Alho and Kangas (1997). The estimated variance can be utilized as an input to the sensitivity analysis.

As discussed earlier, these estimations give provisions for introspection, revision and dialogue as well as estimation of uncertainty to be used in sensitivity analysis.

### 8.4 Verification & Validation

Verification is examined by testing the model against the specifications as well as the data used to construct the model. Validation is always carried out through independent data and in hitherto new conditions. Significant uncertainty in deciding validation is encountered due to limited data, but also the availability of single history while the actual complex system is path dependent and stochastic. We carry out validation through modified Turing’s tests, correspondence testing and model docking. Since our discussion focus is on decision making under uncertainty, we will skip the validation section of the KE process, we go to sensitivity analysis to look at what happens to the uncertainties estimated during other times.

### 8.5 Sensitivity Analysis

#### 8.5.1 Dealing with Uncertainty, Reliability of Information

The uncertainty involved in GSP model building assigning weights could arise from the following situations:

1. Evidence is unavailable, inadequate, or uncertain,
2. Evidence is available, but is contradictory,
3. Evidence is reliable, but variation occurs between subjective assessments by different experts, and
4. Evidence is reliable, but assessor’s weight assignment results in contradictory weight assignment.

Uncertainty arising in the model is tracked throughout the KE Process and the total uncertainty is fed into sensitivity analysis process.
At the level of evidence organization, the subjective estimate of the reliability of the evidence is recorded. Once a piece of information is assessed to be of low reliability, revisiting the sources of evidence and seeking more information is an option, but anything beyond desktop search will be expensive. If information cannot be found by desktop search, commissioning drill down research could be deferred until after carrying out a sensitivity analysis and determining the extent to which the model would depend on the evidence.

The consistency of these evidences with remaining available materials should be used as a check. Simple checks on the information may sometimes help provide an estimate of reliability of the information. In Crusade models, evidence of praise of character or deed by the usual critics and the criticism by the peers were given higher reliability (note: such as rule is context dependent). If authoritative information is available, and that authority could be confirmed, then the material could be used at face value, at least for the initial study. Even when reliable evidence is available, the assessor’s model building is subject to biases, and weight assignment often results in contradictory weight assignment. And, the qualitative materials and expert judgments must be translated into quantitative cognitive and personality structure. In these subjective processes, cognitive biases and human errors could degenerate the quality of the model. The design is intended to minimize these errors, as well as characterize the uncertainties.

Differential diagnosis is not only aimed at arresting cognitive bias of confirming evidence (a significant bias in model construction), but also estimates the value tree level uncertainties taking into account the framework of evidence and their reliabilities. And then during the weight assessment, the intra- and inter- expert variations can also be quantitatively estimated through random variable based regression technique. In addition, the human errors resulting from incorrect assignment can be minimized though the AHP based scoring scheme’s own consistency check.

It must be noted that in all these case, the depth of analysis is kept commensurate with the scope and significance of the assessment.

Sensitivity analysis is used to determine the dependency of the model on the given set of input parameters. Besides, one also needs to explore the decision space to discover strategies. The latter (discovery), being the purpose of the model building, is important. We use both parametric and non-parametric techniques. We omit details of these standard techniques here.
**Personal Information:** My name is Haitao Yin, and I am currently a doctoral candidate at Business and Public Policy Department, the Wharton School of University of Pennsylvania. My area of interest is environmental risk management. If more information needed, please contact me at haitaoy@wharton.upenn.edu or 215-898-6747.

**Supervisor:** Professor Howard Kunreuther is supervising my research on the proposed and other related projects.

**Project Title:** Public Insurance, Private Insurance and Risk Prevention – Has Private Insurance Helped to Reduce Underground Storage Tank Release Risk?

**Project Summary:**

The theoretical literature has discussed the strengths and weaknesses of public insurance and private insurance in financing risk. The advantage of public provision of insurance lies in its ability to generate particular kinds of information, such as the total coverage one buys (Pauly, 2003) or the minimum level of risk prevention investment (Lee, 1992) and so on. However, most agree that private insurance might be a better choice if the risk is identifiable and the total coverage is verifiable. The reason is that the private insurance market determines premiums based on the insured’s own risk and therefore is an effective tool to promote risk reduction behaviors (Freeman and Kunreuther, 1997). On the other hand, if the premium is not tied to risk and the coverage payment is residually financed out of a common pool of revenue, which usually is the case for public insurance, careless behaviors could be encouraged because the insured internalizes only a fraction of the social marginal costs (Persson and Tabellini, 2002). This study aims to provide an empirical investigation of whether private insurance promotes risk reduction efforts more effectively than public insurance, as the theoretical literature has suggested.

The policy variations in financing Underground Storage Tank (UST) release risk provide a great setting for this purpose. In 1988, EPA published regulations that required UST owners and operators to demonstrate that they are financially capable of cleaning up releases and compensating third parties for resulting damages (called Financial Responsibility Requirement – FRR). Among others, Michigan (MI), Illinois (IL) and Indiana (IN) all established state fund programs – typical of public insurance – to help tank owners and operators achieve compliance with FRR. State fund programs simply ascertain if claimants qualify for damages and meet some certain criteria, and then pay their claims. A state fund typically requires all owners to pay equivalent fees. In addition, the taxes used to account for the majority of the revenues of the state funds pool are not linked to up-keep. Because the cost to participate in State Fund program is not related with the release risk, little incentive is created to encourage risk prevention efforts. Michigan discontinued its state fund program in 1995, and required tank owners and operators to obtain private insurance in order to stay in compliance with FRR. As other private insurances, the availability and cost of UST insurance is closely linked to the release risk of each individual tank. Furthermore, a tank release might trigger a premium increase afterwards. Through these mechanisms, private insurance will encourage more risk prevention behaviors and therefore reduce tank release risk. The public-to-private policy change in MI provides us a unique opportunity to examine
whether such policy change has effectively reduced tank release risk, with IL and IN as controls.

The central question of this study is: does the transition from public insurance to private insurance in MI promote risk reduction investment and therefore reduce UST release risk, compared with what would happen if there had been no such policy change. This study is of interest for a few reasons. First, financial difficulties have been plaguing many of the UST state fund programs. The most recent data show that more than 13 states have outstanding claims greater than the current balance. Nationwide, there is a $250 million deficit in total (State Financial Assurance Funds Survey, 2003). This and the fact that about 10,000 new releases are confirmed every year concern policy makers and other stakeholders. There have been discussions about how to reduce tank releases more effectively, how to tackle the huge state fund deficits and whether state fund program should be continued at all. This research will shed light on all these questions and therefore bears important policy significance. Second, the significance of this study is not limited to the UST regulation itself. The choice between private insurance and public insurance in financing risk is not only an issue in environmental context, but also a hotly debated issue in bank deposit, pension benefit etc. (Cooper and Ross 1999). Therefore, this study is not only important for financing environmental risk, but also insightful for other risks with similar characteristics.

The current status of this project is that I have already collected the data and finished preliminary analyses. My findings suggest that compared to public insurance programs, the use of private insurance does encourage risk prevention behavior on the part of tank owners and operators, and therefore reduce UST release risk.

**Detailed Budget: $4270 in total**

- **Collecting Data:** $350
  
  *Note:* The first round of date collection work is already finished. To have a better understanding of data related issues, I need to make phone calls to the persons who collected and managed these datasets. $50 is for these phone calls. I am contacting the Society of Independent Gasoline Marketers of America (SIGMA) to buy their annual statistical reports. $300 is a rough estimation for the cost of the reports (1986-2003).

- **Purchasing a Computer:** $1400
  
  *Note:* My computer only has 512 MB RAM, which is not capable of processing large size of dataset. My three-state (MI, IL and IN) UST database is more than 300 MB. Sometimes it takes ten minutes or more to finish one simple command. Sometimes the computer just stops working.

- **Meeting with Officials at EPA Underground Storage Tank Office (OUST).** $300
  
  *Note:* As I move forward with this project, I expect to meet officials at EPA OUST. This budget will cover two trips to Washington DC.

- **Submitted Proposal for the Poster Sessions of Association for Policy Analysis and Management (APPAM) Conference at Washington DC in Nov. 2005 and Attending this Conference** $700
  
  *Note:* This budget covers one round-trip train ticket ($100), registration fee ($80), three days of hotels ($400) and meals ($120).

- **Submitting Paper for the Poster Sessions of Society for Risk Analysis (SRA) Meeting**
at Orlando, FL in Dec. 2005 and Attending this Meeting $825

*Note:* This budget covers one round-trip air ticket ($400), registration fee ($105), three days of hotels ($200) and meals ($120).

- Attending American Economic Association (AEA) Meeting at Boston, MA in Jan. 2006. $695
  
  *Note:* This budget covers one round-trip air ticket ($300), registration fee ($75), three days of hotels ($200) and meals ($120).

**Current Sources of Research Funding:**

I have no particular research funding. Currently I receive a research assistantship and teaching assistantship from Risk Management and Decision Processes Center of the Wharton School.

**Research Assistantship:** I am currently a Research Assistant for Peter Schmeidler, who is a senior research fellow at Risk Management and Decision Processes Center of the Wharton School. We work on a survey of ISO 14001 certificate holders, which aims to look at the environmental and financial value of ISO 14001 Certification.

- Develop Research Hypotheses;
- Design Questionnaire;
- Perform Data Analysis;
- Write up the Reports;
- Present the Research at National Manufacturing Week Conference in March, 2006

**Teaching Assistantship:** I will be a Teaching Assistant for Prof. Kunreuther for the class of “Risk Analysis and Environmental Management” in Spring, 2006.

- Hold Office Hours;
- Grade Assignments and Examinations.
Proposal for Russell Ackoff Doctoral Student Award
for Research on Human Decision Processes

Title
Envy and Portfolio Allocation in Defined Contribution Pension Plans

Student
Jacqueline M. Volkman, Insurance and Risk Management

Faculty Advisor for Project
Alexander Muermann

Summary of Project

There has been a movement over the last decades in retirement saving plans offered by employers from a defined benefit (DB) to a defined contribution (DC) system. In DB plans, the plan sponsor is liable for pension benefits promised to workers and asset management of such plans; DC plans, on the other hand, allow employees to decide into which assets their retirement investment will be allocated. That is, in DC plans, the plan participant is responsible for asset allocation and management. Recent capital market volatility with the 2000 “tech bust” and 2001 Enron scandal has affected the performance of many plan participants’ portfolios. As more DC plans are offered and workers manage their own retirement savings, it is important to understand how they make investment decisions.

Envy is a behavior that is relevant to decision making in pension plans. Planning for retirement is likely to be influenced by a behavioral element (Zeckhauser, 1986). It has been shown through observed behavior in institutions that individuals planning for retirement discuss their investment strategies with each other (Duflo and Saez, 2003; Thaler and Benartzi, 2004), and that workers follow the investment strategies of their peers (Hong et al, 2004; Patel et al, 1991; Hong et al, 2005). Furthermore, many have argued that individuals’ preferences are dependent, not only on their own wealth, but also on the wealth of other individuals (Veblen, 1899; Easterlin, 1998; Luttmer, 2005). Therefore, it is very likely that, as a worker is aware of his co-worker’s retirement investments, his utility will depend on his co-worker’s DC performance in addition to his own DC performance. Since workers discuss their DC performance with each other and individuals have preferences that depend on the relative position of their own wealth to
that of their co-workers, co-worker envy could have a significant effect on a worker’s portfolio allocation in their DC pension plan.

In this project a theoretical decision model that incorporates individual preferences exhibiting disutility from “envy” will be developed. Such envy can arise, for example, if one’s co-workers make better investment decisions in their pension plans. The effect of anticipated envy on portfolio allocation in DC pension plans will be examined. I will solve the model for Nash equilibria and am interested in determining if envy will lead to a riskier portfolio allocation in equilibrium than that of a purely risk-averse individual.

The model includes a utility function to demonstrate the investor’s preferences, which includes a second attribute for envy that is dependent on the co-worker’s DC portfolio performance. I examine different structures for this second attribute. Prior literature has noted that when modeling such behavioral preferences, one should consider the gains and losses relative to a reference point. In this model the reference point is the co-worker’s DC performance. Additionally, loss aversion implies that an individual’s preferences to be risk-averse over gains but risk-seeking over losses. Structural forms that account for loss aversion are considered.

Whether envy impacts adversely on the portfolio allocation of pension plan participants is also empirically important. Predictions from this theoretical model could be tested from current data sources or through experiments. I plan on testing the predictions from this model empirically in the future.

References


### Proposed Budget

<table>
<thead>
<tr>
<th>Item</th>
<th>Amount</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Risk and Insurance Association (ARIA) Annual Meeting</td>
<td>$800</td>
<td>Washington DC, August 2006 Attendance and Presentation of Paper</td>
</tr>
<tr>
<td>European Group of Risk and Insurance Economics (EGRIE) Meeting</td>
<td>$1500</td>
<td>Barcelona, September 2006 Attendance and Presentation of Paper</td>
</tr>
<tr>
<td>Summer Support</td>
<td>$1400</td>
<td>Produce a Final Copy of the Paper</td>
</tr>
<tr>
<td>Software – Scientific Workplace</td>
<td>$300</td>
<td></td>
</tr>
<tr>
<td>Books</td>
<td>$300</td>
<td></td>
</tr>
</tbody>
</table>

### Other sources of research funding

Shannon Schieber Doctoral Fellowship
Risk Aversion and Loss Aversion in the Executive Control Processes

Jeff Larson

Previous literature has posited a number of decision strategies that humans use when making decisions. Some of these strategies include weighted adding, elimination by aspects, satisficing, etc. None of these techniques is used universally, which implies that a crucial part of every decision is the determination of the strategy to be used to make any decision. Very little literature has examined the governing dynamics of this part of decision making. A notable exception is a literature stream undertaken by John Payne, James Bettman, and Eric Johnson. In this research stream, they posit that the various decision strategies each have a different expected accuracy as well as a different expected effort level. The optimal use of decision strategies depends on two factors: 1) the desired balance between effort and accuracy, and 2) the structure of the decision situation. This is illustrated in Figure 1. In the first frame, strategy A, B, or D is optimal, depending on the desired balance between effort and accuracy. In the second frame, the positions of the various strategies have shifted, due to a change in the structure of the decision problem. For the given effort/accuracy level, B was the optimal strategy in the first frame, but A becomes the optimal strategy in a decision with a different structure.

The literature stream undertaken by Payne, Bettman and Johnson has demonstrated that people are highly adaptive to both factors. The authors showed in a series of experiments that study participants switched between compensatory and noncompensatory processing as called for by the structure of the decision situation as well as by time pressure. The focus of this literature was to demonstrate that the various posited decision strategies all have their merits and are used under various conditions. It posits an executive control process that determines which decision strategy will be optimal. It does little, however, to characterize any properties of this executive control process. Does the executive control process calculate the expected benefit of each strategy and choose the strategy with the highest expected benefit? If so, what is the shape of the value function? I hypothesize that the value function is concave, yielding a risk averse executive control process.

Additionally, if the executive control process calculates the expected benefit of potential strategies, does it accurately assess the efficiency and accuracy of each decision strategy? Judgments over such ambiguous executive control parameters as accuracy and effort are likely subject to a number of biases. The executive control process’ evaluation of accuracy and effort is likely to be very context-dependent. I hypothesize that evaluation of efficiency and accuracy will demonstrate loss aversion. In other words, previous decisions will form a reference point from which future evaluations of efficiency and accuracy will be evaluated as “gains” and “losses”.

Consider a decision maker who makes a given decision with some regularity, such as the choice of jam, soup, milk, laundry detergent, etc. Consider further that past experience has shown that an elimination by aspects decision strategy optimally balances this decision maker’s desire for accuracy as well as the desire to conserve mental energy. This is akin to the first frame of Figure 1, where B represents an EBA strategy. Now consider that the decision environment changes such that an EBA strategy is no longer optimal, as in the second frame of Figure 1. Now, strategy A, a satisficing strategy, is optimal. The executive control processes face a risky proposition—continue to use the non-optimal strategy, which is known to provide a given level of effort and accuracy, or
seek to find a more optimal strategy, which provides a better balance of effort and accuracy, but which will be found with probability less than 1. By hypothesis, the executive control processes will exhibit risk aversion in such a situation. Thus, when decision makers learn to use a given strategy in a decision environment, they will continue to use that strategy even when the environment changes such that the chosen strategy is no longer optimal. For this reason, they will perform more poorly even than decision makers who face that decision environment for the first time.

In addition to risk aversion, the executive control processes are also hypothesized to exhibit loss aversion. Consider a decision maker who faces a new decision environment with a given desire for expediency (low effort) and accuracy balance. After facing this decision environment several times, he faces the decision environment under greater time pressure, altering the desired balance between expediency and accuracy. The previous balance between expediency and accuracy is hypothesized to form a reference point from which changes are evaluated as gains and losses. The need for greater expediency must by nature come at a loss to accuracy. Since losses are more aversive than equivalent gains are rewarding, the decision maker will not adequately adjust the expediency of his decision, but rather remain closer to the original expediency-accuracy point.

Two experiments have already demonstrated both hypothesized effects, but more experimentation is needed to rule out competing explanations. In one experiment, participants were required to learn a simple associative rule. Those who began the task under high time pressure (manipulated through the reward structure) were less likely to learn the associative rule than those we began the task under low time pressure. In other words, once they had settled on an efficient strategy, they were less likely to attempt to discover the more accurate strategy. I have already planned a follow-up experiment that will better demonstrate the effect by allowing for better tracking of the particular strategies being utilized.

Another experiment has demonstrated loss aversion in the executive control processes. In that experiment, participants examined two lists of numbers to judge which of the two lists had the higher sum. Unlike the previous experiment, this one did not examine different decision strategies. Instead, participants merely had to determine how long to examine the two lists before making their guess. Very few participants actually summed the numbers. Those participants who started under high time pressure continued to make their guesses more quickly and less accurately when time pressure was removed that those who started without time pressure. Similarly, those participants who started without time pressure continued to make guesses more slowly than those who started with pressure. Further experiments will be conducted to ensure that participants did not alter strategy, but indeed only altered the amount of time they examined the numbers to determine the correct answer.
Faculty Sponsor: Robert Meyer

Budget:
Experiments: $2500
   Participants’ pay will be determined by their performance in the experiment.
Travel: $1500
   Airfare, hotel, conference fee, and incidentals for ACR 2005 in San Antonio, TX
   Airfare, hotel, conference fee, and incidentals for SJDM 2005 in Toronto, ON

Current funding resources:
   Robert Meyer’s research budget.
The Psychology of Two-Part Tariffs

Joseph P. Redden *

September 2005

* Joseph P. Redden is a doctoral student (jredden@wharton.upenn.edu, phone: 215.573.0540, fax: 215.898.2534) Wharton School, University of Pennsylvania, 700 Jon M. Huntsman Hall, 3730 Walnut St., Philadelphia, PA 19035

This work is being completed with Stephen J. Hoch, also of the Wharton School.
RESEARCH DESCRIPTION

Marketers frequently use nonlinear pricing plans, where the mapping from quantity purchased to total price is not a strictly linear function. An example is a two-part tariff which consists of an upfront flat fee plus a per unit surcharge. Traditional economic analyses of nonlinear pricing plans have assumed that the consumer calculates an accurate and unbiased effective cost and has no preferences for specific attributes of a pricing plan beyond its expected cost (Dolan 1987; Kohli and Park 1989; Weng 1995). This research project questions these assumptions and proposes that consumers exhibit systematic tendencies when choosing a nonlinear pricing plan.

The effective cost of a nonlinear pricing plan is not always readily apparent. For example, the reader should attempt to estimate the effective cost per unit of the following plan if the expected usage is 500 units: $28 for 350 units and $0.22 for each additional unit. Although estimating the effective cost per unit presents a cognitive challenge, the formation of preferences may be less difficult since it requires only relative judgments on each attribute. The attributes for a two-part tariff include the overage rate, the dollar amount of the flat fee, and the number of units included in the flat fee (i.e., the only three variables in a two-part tariff). Consumers can compare pricing plans on these attributes using a simple set of comparison rules (Alba and Marmorstein 1987; Russo and Dosher 1983; Weber, Goldstein, and Barlas 1995). Specifically, these rules should favor lower rates for additional usage, the inclusion of more units with the flat fee, and smaller dollar flat fees. If people use these rules, they will prefer pricing plans with comparative advantages on a greater number of attributes beyond any differences in cost. Further, the presence of uncertainty about expected usage encourages the use of simplifying heuristics.
by increasing the difficulty of any calculation efforts. Thus, people should use these
simple attribute comparisons more with increased uncertainty about usage.

Although consumers might compare pricing plans on each of the attributes
equally, previous research shows that consumers generally exhibit loss aversion
(Kahneman and Tversky 1979). If consumers have the goal of limiting their exposure to
the losses from high overage costs, they will favor plans that include a great number of
units and have a low overage rate at the expense of a higher dollar flat fee. Here, a higher
flat fee is the cost of insurance to protect against the unknown and potentially unlimited
overage fees. Therefore, consumers will tend to give more weight to the attribute
comparisons for the overage rate and number of units included in the flat fee. As
previously predicted, this tendency will only increase with usage uncertainty.

**PROGRESS TO DATE**

A study tested these predictions by asking 121 participants to choose between two
different two-part tariff pricing plans for an undefined product. Each participant saw their
expected usage level and the terms of each pricing plan. As well, some participants were
provided the minimum and maximum usage level to increase the salience of usage
uncertainty. The results confirmed the theoretical predictions. Specifically, after
controlling for differences in actual costs, people preferred pricing plans with lower
overage fees and plans where the flat fee included a larger number of units. However,
people did not display a preference for plans with a smaller flat fee. These preferences are
attributed to a desire to limit the potentially unlimited upside loss/risk associated with
much higher than expected usage. Further, when the range of expected usage was made
salient, choices reflected an even greater focus on limiting potential losses through lower
overage rates and flat fees that include more units.
FUTURE DEVELOPMENT

Although the previous study supported the theoretical predictions, it did not link these preferences to a conscious goal of avoiding excessive losses. Future studies will demonstrate that people have the goal of limiting losses by default, but can have other salient goals as well. For example, if a consumer wants to limit their commitment, they would tend to choose the plan that minimizes the size of the flat fee. Similarly, consumers who have recently had a great deal of overage charges should pay more attention to the overage rate. More generally, the importance of each pricing plan attribute will depend on the salient goals at the time of choice. Consumers and marketers should understand the role of goals in shaping consumer choices of pricing plans.

The previous study chose an undefined product context to control each participant’s information about their usage. In actual settings, consumers will often need to construct a distribution of their expected usage from memory. However, previous research has not explored how consumers construct this distribution or apply it to their choices. For example, consumers may base their choices on a single usage level that is particularly accessible (e.g., usage for peak period or last period). Likewise, consumers may evaluate pricing plans using several different usage levels. Here, consumers may choose the plan that compares favorably on more attributes for more of the sampled usage levels. Consumers could benefit from knowing how to estimate their usage and evaluate pricing plans to pick more satisfying pricing plans.

PLANNED USE OF FUNDS

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral Experiments</td>
<td>4 studies x 100 subjects/study x $5/subject</td>
<td>$2,000</td>
</tr>
<tr>
<td>Conferences</td>
<td>2 academic conferences</td>
<td>$1,500</td>
</tr>
<tr>
<td>Practitioner Study</td>
<td>Test with Bell South business customers</td>
<td>$500</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>$4,000</strong></td>
</tr>
</tbody>
</table>

Sole current source of funding is research budget of Stephen Hoch (my advisor)
REFERENCES


Trust is a basic precondition to a social contract. Without it, people are severely limited in their ability to carry out economic transactions, and norms of cooperation are nearly impossible to realize. Though there has been research on trust in many fields a number of open questions remain.

Economists have done empirical work on bilateral trust using the Trust Game, but it has yet to be established whether trusting behavior is normative. Ought we trust in situations in which trusting behaviors are risky? How do we test this hypothesis? One potentially fruitful method of analysis is examining what happens when trust is violated. To what extent are those whose trust has been violated willing to pay to punish the transgression? Does this willingness hold constant across multiple risk levels, or are people sensitive to the strains of commitment as risk is increased for the other person?

The literature also lacks an examination of whether groups are more or less trusting than individuals. As has been shown in Commons Dilemma experiments, individuals in groups often exhibit free-riding behavior. In group trust situations, such as N-person Prisoner's Dilemmas or Stag Hunts, what are behaviors when the norm is made salient, either by allowing for non-binding commitments to act, or by offering costly options to punish non-cooperators? These results would be an interesting foil against bilateral trust results. Are the strains of commitment similar in group contexts, especially if threshold effects are introduced? By examining the punishment response against those who violate the group trust, we can see how sensitive people's moral sentiments are to differing conditions.

Perhaps most salient to the predictive power of any theory of trust would be the cultural differences in trusting behaviors. One potentially fruitful indicator of variance in trusting behaviors would be variations in social structures. Differing social structures can accommodate varying levels of trust and trustworthiness. A unique case study of this phenomenon is Italy. Northern and Southern Italy have very different social structures, and are anecdotally documented to have significantly differing levels of trust of newcomers (Putnam, Robert D., Robert Leonardi and Rffaella Y. Nanetti Making Democracy Work. Princeton: Princeton University Press, 1993. ). A study that compared the trusting behaviors of northern and southern Italians would provide an interesting test of this phenomenon. The north-south difference in trusting behaviors has never been experimentally tested, but the result has been widely assumed. Any results would be of immediate interest to the wider academic community. Dr. Bicchieri and I have had preliminary discussions with colleagues at Trento University about collaborating on this project.

Experimental results from the studies sketched above will provide ample data to inform a more complete theory of trust, and should inform future research programs. This will hopefully lead to both a more fulsome philosophical understanding of our moral sentiments, but also provide the data needed to construct models of group trust under
uncertainty that might offer some predictive capability. Especially with an enhanced understanding of the role of social structure, computer models could be developed to examine the different resilience levels of varying social networks to untrustworthy individuals. This could be extended beyond just commons problems to signaling games more generally, which has significant bearing on Interdependent Security, gossip networks, and other information-sensitive fields.

The budget for this project is as follows:

**Travel Expenses: $3000**
- Plane ticket to Italy is between $1000 and $1400
- Remainder to be used to offset food and lodging expenses

**Experiments: $6000**
- 200 study participants

Currently the only other source of funding for this research is my doctoral stipend. I am actively seeking out additional funding sources.
Service with a Snub: Using Shame to Gain in the American Wedding Dress Industry

Student: Lakshmi Ramarajan

Faculty Advisor: Sigal Barsade, Management Department

Descriptive Summary

Service with a smile has long been the approach of customer service, especially in retail. The basic mechanisms are those of emotional contagion where the service person’s positive emotion generates a positive emotion in the customer which induces buying or reciprocity in which the social exchange element of positive behavior results in a sale (Grandey et al, 2005; Pugh, 2001; Sutton & Rafaeli, 1988; Rafaeli and Sutton, 1987). However, this paper proposes that the service-with-a-smile model is not generalizable. For example, emotions may differ depending on the industry. In the insurance sales market, salespersons’ may manipulate fear and anxiety in the potential customer in order to sell insurance. The particular strategy or niche market that the firm executes or occupies could also have an influence on the emotional strategy used to sell, for example, a high-end retail store that pursues a differentiation strategy may use shaming (of customers) or extreme pride (of employees) to emphasize their distinctiveness.

The use of shame in a particular segment of the wedding dress industry is used as an example of this phenomenon. The decision to purchase a wedding dress - what kind, where, how much to spend - is typically invested with the weight of a bride’s dreams. The wedding dress is a symbol of the bride’s ideas of who she is (Clarke, 2005; Walker, 2000; Wolf, 1995). However, this decision is also fraught with uncertainty of two kinds. The bride, in most situations, is a novice to the wedding industry. Despite the fact that this is such a highly significant purchase, many brides have little to no experience or expertise in acquiring wedding dresses. In addition, in many respects, wedding dresses cost more than the average purchase many women are used to making (Clarke, 2005). Thus, by entering the market for wedding dresses, brides are moving between class and status-levels, often willing to spend more than they normally would on an item of clothing. Both the lack of expertise or information and the status move can contribute to high uncertainty. However, the decision itself is invested with high importance. The basic hypothesis is that in such situations, a shaming approach to the customer service interaction rather than the service-with-a-smile model can result in higher gains for the seller.

Shame is defined as a response to a personal failure attributed to the self that arises from a concern with other’s evaluations that the individual has not lived up to others’ expectations (Bagozzi, Verbeke & Gavino, 2003; Ashforth & Humphrey, 1995; Fischer & Tangney, 1995) and it is generally associated with a need for others’ approval (Baumeister, 1995). Thus, when an individual is shamed, they are made more self-aware of their own limitations in a particular social context. In the case of the wedding dress transaction salesperson’s use of shame in the encounter should trigger or accentuate both the uncertainty associated with the status move and the information uncertainty inherent in the situation for the individual. The salesperson themselves probably do not engage in shaming behavior unless they themselves observe status cues which allow them to tailor their treatment of the customer. For example, Mattila and Enz (2002) discuss the
personalization of the service encounter, such as observing transaction-specific cues to adjust the customer service encounter accordingly. This observation and adjustment need not be explicit, it could be a highly implicit process, such as stereotyping, or picking up on customer displays of emotion or behavior, such as anxiety or nervousness in an uncertain situation. Once the shaming process has occurred, the customer is likely to try to resolve either the status uncertainty or the information uncertainty. Research on uncertainty and judgment/decision making tells us that in highly uncertain situations, individuals seek diagnostic cues to guide their behavior (Kahneman, Slovic & Tversky, 1982; Weary & Jacobson, 1997). Moreover, in the social cognition literature, person perception research shows that in uncertain situations individuals rely on categories and stereotypes to make sense of the social world (Macrae & Bodenhausen, 2001; Molden & Higgins, 2004). In this case, the salesperson’s authority which rests on both their status and expertise provides cues which may result in the customer’s reliance on the salesperson’s guidance as the authority figure, and may result in their paying more. The customer may also try to make an expensive purchase in order to restore their self-esteem after the shaming process, which will also result in their paying more.

Proposed Method
In order to test the hypothesis that customers purchasing an item which is highly relevant to them but about which they have a) little prior experience and b) is something out of their normal purchasing price range and to see if therefore they are willing to accept the salesperson’s authority and pay more for their goods due to shaming rather than smiling, we propose two studies:

Study 1: Lab experiment simulating the customer service interaction

Procedure
A between-subjects lab experiment is proposed in which participants would be randomly assigned to one of three conditions, Shame, Neutral or Smiling. In each condition, participants will be told that they are going on their first major job interview with a prestigious firm and they do not have an interview suit. They will also be told that the interview suit is important to making a good first impression. Finally, they will be told that they have no experience buying an interview suit and that their normal clothing purchases are in the range of $50-$100. They will also be informed that the suit will be one of the most expensive things they will have ever bought. The participants will be told that they have a budget for the suit and if they exceed their budget they will be forced to cut back on other important expenses. In addition, to make the cost real and engage the participants in the experiment, there will be a graduated payment scheme from $5-$15.

When they enter the lab the participants will be told that there are several suits from which they can choose. To control for actual differences in quality, we will provide three suits which we would have purchased in the same price range (of slightly different styles but in the same color) but change the pricing information on the tags. The salesperson will be a trained actor acting as a confederate and instructed to subject the participant to shaming, smiling or affectively neutral behavior. The interaction will also be recorded to provide an objective measure of the emotion induced and experienced. The actual price of the suit chosen will be the main dependent variable. After the experiment, participants
will be given a post-hoc questionnaire to assess whether they really felt shame, nothing or positive emotion and then debriefed.

**Measures**

**Independent Variables:**
Shame: Herrald & Tomaka (2002) use a shame manipulation based on Core Relational Themes (CRTs) (Lazarus, 1991) in which the confederate conveys verbal and non-verbal cues indicating failure to live up to personal standards. Example phrases include: “I think you can do better” or “I don’t think you are doing as well as you can.” These can be adapted for the customer service context.

Smiling: Grandey et al (2005) used a trained actress as a confederate to display both authentic and false smiles in a customer service interaction. To induce an authentic display of positive emotion the actress was told to think of the incoming customer as a chance to make someone feel good and encouraged to use her acting skills to display genuine positive emotion.

Neutral: In the affectively neutral condition the salesperson would simply tell the customer to take a look at what they like and bring it to them when they are ready, but would be instructed not to interact with them in any other way.

**Manipulation check:**

To serve as a manipulation check for the emotions exchanged in the interaction, each customer service encounter will be videotaped. Expert coders will then view each encounter and classify whether they observed treatment and responses that coincided with shame, neutral or smiling. The videotaped interactions could also be used to rate the strength of the emotion observed in the encounter and whether there were other emotional responses by the customers in addition to the dominant emotion that was induced.

**Dependent Variables:**

Price of the Suit: The main dependent variable will be the price of the suit the participant chose. As we are already controlling for quality of the suit in the experiment, the variation in price should be attributed to the emotion present in the customer service encounter.

In addition, a post-hoc questionnaire would also assess the extent of shame, positive emotion or general mood (in the neutral condition) generated by the encounter.

Shame: The Bagozzi et al (2003) measure of the experience of shame could be adapted for the customer context. This measure has 15 items assessing self-focused cognition and physiological changes related to the experience of shame. Sample items would include: “I think the salesperson considers me a failure as an individual” and “I feel like shrinking.”
Smiling: Grandey et al (2005) asked participants to indicate whether the following positive behaviors were displayed (smiling, making eye contact and speaking in a rhythmic tone. In addition items used to check the authenticity of the displays would also be included, such as, “the salesperson is faking how s/he feels in this interaction,” and “the salesperson is pretending, or putting on an act, in this interaction.” Finally, items to determine self-ratings of happiness of the customer will also be included.

Neutral: In the neutral category, individuals will only be asked to complete an adjective-based survey to determine the generic mood of the individual after the encounter. Subjects in all conditions, including neutral, will be asked to complete the The Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988) which has been used in prior research to measure generic positive and negative affect.

Customer Satisfaction: Participants will be asked “To what extent do you feel satisfied with the encounter” and “to what extent do you think you will return to this store/buy from this salesperson again”

Control variables: Sex-match and friendliness of the salesperson could be used as control variables. Friendliness can be measured with a scale from Tsai and Huang (2002) that includes items such as “provided the service in a friendly manner,” “had a kind smile,” and “treated the customer nicely.”

Study 2: Survey of Women Purchasing Wedding Dresses

The second study would be a web-based survey accessible through some wedding website, such as theknot.com or weddingchannel.com for prospective brides who have bought their wedding dresses in stores.

Independent Variables:
Emotion: Individuals will then be asked to describe in their own words their customer service encounter. Independent coders will code this text for descriptions that align with the experience and behaviors related to shame or smiling. Customers will also be provided with the same questionnaire used in the above study to determine their experience after the encounter.

Dependent Variables:
Price: Participants will be asked to name the price they paid for the dress

Customer Satisfaction: The customer satisfaction questions would be the same as above and also include “to what extent will you recommend the store/salesperson to a friend or family member buying a wedding dress.”

Control variables: Individuals will be asked for demographic data, including the sex of the salesperson, their average income, zip code and average spending on clothing, as well as the zip code of the store location where they purchased their dress. The zip codes will be used to furnish status/class differences between the customers normal purchasing and living experience and that of the wedding dress store. Participants will also be asked the
brand name of the dress they purchased in order to control for quality differences between different ranks of designers (e.g. Melissa Sweet vs. Vera Wang).

<table>
<thead>
<tr>
<th>Detailed Budget</th>
<th>Number of People</th>
<th>$/Hour</th>
<th>Number of Hours</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confederate</td>
<td>1</td>
<td>20</td>
<td>100</td>
<td>2000</td>
</tr>
<tr>
<td>Participants (study 1)</td>
<td>90</td>
<td>10</td>
<td>1</td>
<td>900</td>
</tr>
<tr>
<td>Participants (study 2)</td>
<td>100</td>
<td>5</td>
<td>1</td>
<td>500</td>
</tr>
<tr>
<td>Video Technician</td>
<td>1</td>
<td>10</td>
<td>90</td>
<td>900</td>
</tr>
<tr>
<td>Total Amount Requested</td>
<td></td>
<td></td>
<td></td>
<td>4300</td>
</tr>
</tbody>
</table>

**Budget Notes:**

Confederate: A trained graduate acting student will be required to ensure the plausibility of Study 1 and induce the appropriate emotion in the participants. We have requested a 100 hours to include both the actual experiment and the training time needed to prepare the actor.

Participants: At least 90 students will be needed for study 1, i.e., 30 students per condition. An average of $10/participants is requested for Study 1. In addition, approximately 100 respondents will be needed in study 2 to ensure sufficient numbers for both shame and positive emotion experiences. Approximately 5$/participant is requested for Study 2.

Video Technician: Finally, support for a video technician is requested for approximately 90 hours to tape the interactions in Study 1.

**Other Sources of Funding:**
We have approximately $3000 from other sources of funding.

Prof. Sigal Barsade will provide $1200 for two RAs for data entry, coding and video rating.

Local Clothing Retailer will provide $1800 in-kind contribution of 6 suits. Three men’s suits and three women’s suits will be needed for the experiment. Given we need to control for quality/price and color differences, it is likely they will have to be purchased new. We propose to ask local clothing retailers to contribute suits in-kind.
References


Communities of Innovation: Exploring New Knowledge Creation and Evolution

Research Proposal for Russell Ackoff Doctoral Student Awards for Research on Human Decision Processes

Phin Upham (PhD Student, Wharton School, Management)

Project sponsored by Lori Rosenkopf (Management) and Lyle Ungar (Engineering)

Research Question and Summary:

I believe this project furthers the vision of Dr. Ackoff by providing insight into a topic close to his heart – how new ways of innovative thinking change the world and how systems of thought evolve over time. It does so in a context of group innovation in medical and technological innovation. I would be honored to present my work at the conferences planned in Fall 2005 or 2006.

Specifically, I am working over this summer on a cross-disciplinary project with advice and support from Associate Professor Lyle Ungar of the Engineering School and Associate Professor Lori Rosenkopf of the Wharton Management Department and propose to do research on and develop a specialized analytic framework and analysis tool for scientific and medical articles which can be used to identify and predict potential emerging research trends, as well as to predict and track the emergence of distinct and evolving knowledge ‘schools of thought.’ Within the framework of evolving knowledge researcher’s act under joint uncertainty as to which areas are promising and which are developing quickly. By providing a quantitative analysis to this area we hope to shed some light on how groups interact under such conditions of incomplete knowledge to advance science and medicine.

We hope that at least one research paper and a best-of-kind program will emerge from this project. Traditional bibliometric analysis methods will be enhanced using recent advances in clustering technology to allow analysis of the impact of different types of articles and the evolution of research communities. Our main area of study will be the medical literature, but preliminary studies will also be done on the management and computer science literature. We will use a combination of data provided by Thompson ISI, which has decades of experience building databases of abstracts, authors and citations, and similar information extracted by crawling the HighWire collection (with permission), and the web. For the later data, we will use the Rexa software for and extracting authors, their institutions, and citations. Our work is unique in combining large scale information extraction, clustering, and analysis methods with analytic methodologies that have been created to study how scientists interact and how research fields develop.

Theoretical Considerations and Methodologies:

The last decade has also seen major improvements in the speed of clustering algorithms (Ben-David; A. McCallum 2000; S. Guha 2000; A. Popescul 2003). In Management there has been substantive theory work on how fields develop new ideas, how scientists interact, and how new knowledge is created. For example, creators of new knowledge have an incentive to position themselves as a part of a school of thought in order to maximize their impact (Birnbaum 1981; Stigler, Stigler et al. 1995). Birnbaum’s (Birnbaum 1981) work on the subject shows that interdisciplinary studies are harder to publish and take longer to complete. Furthermore, while overly myopic research might not be beneficial to one’s intellectual development (Rosenkopf and Nerkar 2001; Katila and Ahuja 2002; Nerkar 2003) the literature indicates that knowledge progresses as a dialectic between schools of thought and that positioning oneself in such an ‘invisible college’ is advantageous for future impact. But researchers do not and cannot know what area of research will “pan out” and which area is likely to achieve promising results – they act with incomplete knowledge and joint-uncertainty.

However, the lack of appropriate data and the shortcomings of current methodology have limited the field’s ability to test and fine-tune its theory. One of the central aims of the project is to use sophisticated modeling techniques to look at how the medical field grows and develops over time. We are exploring new medical knowledge creation to examine how the choice of sources used to develop new knowledge affects its future use and its success. Moving from medical innovation at the paper level to the group level of
medical researcher in ‘communities of innovation’ or ‘schools of thought,’ we will examine these clusters over time as they evolve and ask what characteristics of in/out cluster and in/out field lead to growth in the cluster in the next few years. By aggregating the ideas on the cluster level we may be able to see beyond the uncertain individual drivers of success and begin to see structural elements of success.

Past methodology in management addressing these analytic issues has failed to sufficiently take into consideration the dynamic and complex structure of scientific knowledge. Using a more powerful clustering method will allow many hypotheses can be tested. Example hypotheses which have been proposed and only partially tested are:

- New knowledge has more impact if it is a part of a school of thought than if it is not.
- A position within the intellectual core of a school of thought tends to impact that school of thought more, while a position on the intellectual periphery of a school of thought tends to produce knowledge with more impact outside its school of thought.
- Knowledge drawn from within a school of thought tends to lead to shorter term impact, while knowledge drawn from outside school of thought tends to lead to longer term impact.

To pursue this path we have hired a computer science PhD student (Peng Bi), who’s work will be supervised by me, Phin Upham, and Lyle Ungar, Assoc. Prof. of Computer Science (Engineering School), who has agreed to oversee in technical aspects of the project such as debugging, coding structure, etc. This student would help collect and process the documents, get existing code (including that from Rexa, a successor to CiteSeer) up and running, develop prototype new analysis code, and produce preliminary results. Although this sounds ambitious, we have good evidence based on Lyle Ungar’s prior work with CiteSeer and current work proceeding in parallel at the University of Massachusetts, Amherst that we will be able to obtain interesting initial results and produce compelling demonstrations. Lori Rosenkopf is working with me in the analytic aspects of this project developing a framework to understand and analyze the evolution and development of technical and pragmatic “schools of thought.” Wharton has agreed to provide major disk space and compute time on a powerful new Linux cluster, so no additional hardware is required.

Research Goals

Our goals are twofold: (1) to develop tools to support the analysis of the development of scientific fields under the conditions of incomplete information and joint uncertainty about technology. This would go toward deepening the understanding of emerging scientific knowledge and begin posing questions about how such knowledge evolves dynamically over time. (2) To write papers building an analytic framework to provide insight into the development and evolution of scientific and medical knowledge development. On the quantitative side, we plan, eventually, to index all of the articles whose abstracts are included in PubMed along with related patents, grant proposals, and additional data available from the internet. The core search and analysis tool would provide: keyword search over the abstracts and (eventually) full-text of millions of scientific research papers, citation linking so the user can follow cross-references---finding not only the papers cited by a given article, but also the papers citing a given article (the "backward pointers"), and links, cross-references and relations not only of research papers, but people, universities, companies, conferences, journals, and grants.

Beyond the research paper we see emerging from this project, the technology we build, among other things, will allow researchers to decrease their incomplete information and see which people have published together, which institutions are most active on a given topic, which papers were most influential, which researchers might be suggested as the best reviewers of a given paper submission, etc. By building this framework around the analytic power of management theory and its insights on the development of fields under conditions of uncertainty and incomplete information, and the emergence of potential new fields, we will create a powerful tool that will yield new insights into medical and scientific research.
**Budget:**

Funds: Total Request $5,000

- Some funds to help cover programming costs (see Note below) = $3,000
- We request $1,000 in travel money to be used to defray costs for visits to Washington to present our initial results to potentially interested parties there, and to California to HighWire and, if needed, the internet archive (www.archive.org) to facilitate data transfer.
- We request $1,000 in funding expected to be required to compensate HighWire for labor and data extraction work done for our project. = $1,000

**Phin Upham** is a Doctoral Student in the Management Department working on his dissertation. He holds a BA with honors from Harvard University. His interests are in the study of schools of thought, paradigm development, the evolution of technical, scientific and organizational knowledge, and in the social construction of meaning.

**Note:** We have received 10,000$ from Wharton’s MACK Center as a summer stipend and for some programming and travel related to this project. We expect this to be sufficient for summer living expenses through September and then additional travel and RA funds to be necessary. We will have exhausted the funds allotted for programming and travel from this grant as of August 15th.

**References**


