Boundedly Rational and Emotional Agents: Cooperation, Trust and Rumor

(rough draft: 4/24/01)

M. J. Prietula
Goizueta Business School
Emory University
Atlanta GA 30322
mike_prietula@bus.emory.edu

K. M. Carley
Social and Decision Sciences
Carnegie Mellon University
Pittsburgh PA 1521
Kathleen.Carley@CS.CMU.EDU

Abstract

Computer-based agents, in various forms, are becoming actively involved in our personal and professional decisions and deliberations. We interact with them; they interact with each other. In this paper we describe a broad Model Social Agent study where we explore how a process model of boundedly-rational agents with emotion behave across increasingly social contexts, and the impact of cooperation, trust, rumor, and deception within those contexts. We report the results of an initial set of small simulations.

Introduction

Our work focuses on understanding the relationship among humans, agents, tasks, and the social situations in which they are engaged. From this, we seek to establish the elemental basis of social behavior and group phenomena, and make predictions about them. Our guide for this effort is the Model Social Agent matrix.

The Model Social Agent matrix (Carley & Newell 1994) provides a two-dimensional categorization scheme that specifies the kind of *knowledge* required by the agent(s) (in terms of increasingly complex social situations), and the kind of information *processing* capabilities required by the agent(s) to operate with that knowledge, in order to exhibit various kinds of individual and collective phenomena. The scheme is summarized (in slightly modified form as will be discussed) in Figure 1.

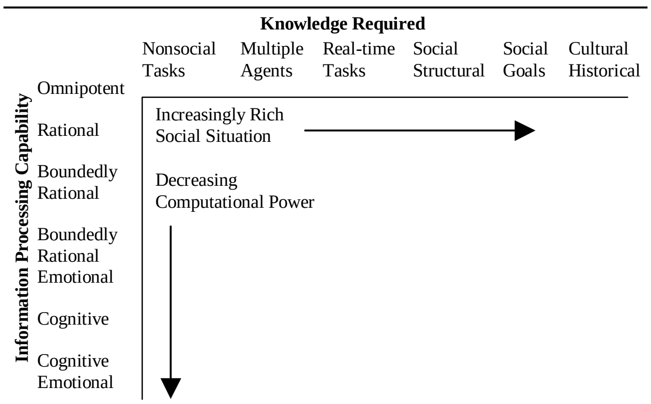

Figure 1. Model Social Agent Matrix

Agent information processing capability range from an agent that is omnipotent as Laplace's demon (i.e., can process all knowledge relevant to the specific situation), to an agent that is both cognitively and emotionally defined (and thus constrained). The categories for this dimension are based on general differing critical assumptions of agent models and would certainly allow subtypes and variations. All agents have the capability to define and alter goals, and interact with other agents and objects. What differs is their capability to exploit the knowledge available within the processing (and sometimes emotional) constraints available. As agents' processing capacity diminishes, different, and perhaps more complex, behavior emerges.[1]

The type of knowledge an agent has regarding social situations range from none (i.e., goals are defined solely in terms of self, any other agents are viewed as objects), to an agent that has a historical perspective and influence of culture, defining beliefs, norms and values. As the social situa-

[1] This is a modified version where we have included Boundedly Rational Emotional Agents – the focus of this research. Note that Cognitive and Cognitive Emotional agents subsume bounded rationality. The strong form of bounded rationality was originally conceived to characterize the effects of a restricted rational agent on the assumptions and conclusions of economic and administrative theory (Simon 1976a, influenced by research from sociology, social-psychology, and cognitive psychology. One component of bounded rationality was that the agent itself was knowledge and computationally restricted (Simon 1955), while another component argued that the environment and its interactions with the agent was an additional source of agent constraints (Simon 1956). Bounded rationality ranges from strong forms that hypothesize the underlying cognitive (e.g., Newell 1990) or institutional (Cyert and March 1963) apparatus to less restrictive derivatives that address various organizational and economic issues at micro and macro levels, such as choice behaviors (March 1978), organizational constructs and market forms (Simon 1991, Williamson 1975), negotiations (Hyder, Prietula & Weingart in press), and relations to complexity theory and economic phenomena (Albin 1998). Overall, bounded rationality (in the strong sense) refers to a *procedural* rationality, that is, to the mechanisms of action and not the outcomes of action, at the individual or institutional level (Rowe 1989, Simon 1976b).

tions an agent is presented with becomes increasingly complex, the richer (and more detailed) the knowledge has to be for the agent to successfully operate in the setting. Consequently, the behavioral repertoire increases as the complexity of the social situation increases. For example, a critical change occurs within the Social Structural setting. Prior situations require that an agent only needs self-defined (e.g., preservation, accomplishment-based) and task-defined (e.g., acquire item) goals. It is here when the agent is faced with goals that are externally (i.e., socially) defined.

The type of agent we are at investigating in this work is the Boundedly Rational Emotional agent. These agents are restricted in both their processing capacity (what they can know) and how they use it. They rely on internal (perhaps flawed) representations of the social and task situations, but are goal-driven and bring to bear knowledge in service of those goals. Our goal for the project is to systematically simulation these agents throughout the range of social situations.[2]

In this work we explore the impact of emotional components on (boundedly rational) agents in a social setting. Emotions serve as a nontrivial component of certain problem solving and decision making processes (Simon 1967). As one might well imagine, there are a host of definitions for "emotion" with recent attempts at incorporating various types of emotion in computational agents (e.g., Bates 1994, Morignot & Hayes-Roth 1006, Picard 1997), with simulations even approximating physiological components and influences (e.g., Canamero 1997).

However, we elect to incorporate a more symbolic form that can be situated in a broader context of problem solving and deliberation in specific settings. Specifically, we focus on the cognitive influences of agent affect, emotion, and behavior. Our view of *emotion* is most similar to that articulated by Ortony, Clore and Collins (1988) as "valenced reactions to events, agents, or objects, with their particular nature being determined by the way in which the eliciting situation is construed (p. 13)." In fact, their theory forms the basis for the elicitation structure of emotion in our model, so we refer to it as the ET component. Under certain conditions, Agents can have emotional responses to events. Basic affective reactions are differentiated with respect to cognitive constraints (as conditions) defining a fundamental set of emotional types.
In our model, however, we must account for four kinds of additional phenomena not addressed by their theory. First, the ET model provides guides for the elicitation functions under a general cognitive model, but we must incorporate some qualitative or quantitative representations that provide the specification of the underlying scales onto which differential values can be mapped. The ET model assures a categorical mapping (intensity variables notwithstanding) into general states, but does not define sufficient information that may carry through to a behavioral mapping of states to events.

Second, we must handle *multiple* events occurring over time. The Agents in our model define a form of repetitive game with memory and consequence. Agents experience sequences of events (some originating from interacting with the world, some originating from Agent communications) that may demand multiple (repeated and perhaps diverse) emotional responses. Agents alter their affective views of the world – they adapt. Thus, a theoretical elucidation of "emotional change and integration" over time must be made. What is the effect of multiple events on the eliciting conditions and subsequent emotional types triggered?

Third, a substantive result of an *emotional* response (as defined by the eliciting conditions) is often a subsequent *behavioral* response to the eliciting conditions (in the contexts of those conditions and emotional state). As emotional states are influenced by events, they too influence event choices by the Agent. Given the situation and emotional state of an Agent, we must define the likelihood of specific subsequent behaviors, and the conditions under which they may arise. The presumption is that these types of behaviors are somehow different from those elicited in a non-emotional state, thus accounting for discontinuities and nonlinear linkages between events and behaviors.

To address these first three phenomena (scaling, multiple events, behavioral responses) we turn from a cognitive model to a model from sociology – affect control theory. Affect Control Theory (Heise 1979, 1987) provides the theoretical substrate and mathematical articulations of how events, agents and objects are perceived (socially) and how those perceptions influence and are influenced by social interactions (or descriptions of interactions). Furthermore, Affect Control Theory affords the mechanisms to predict affective changes (adaptation) and behavioral responses (activity) to such interactions.

Finally, the last phenomenon we explore is the *collective effect* of the previous three phenomena over time, and across multiple agents. Though informal considerations of "group emotions" have often been discussed, little empirical or theoretical guidance is provided for rigorously defining and predicting collective emotional states and behavior as we are considering them. To this end, we are striving to define such models, metrics and theoretical apparatus that can subsume the set of phenomena and explanations woven in the models we are crafting and the theoretical stance we

[2] Agent-based modeling concerns the specification of agent properties and various communication and control constraints. It has a rich, cross disciplinary lineage with early influences from cellular automata (von Neumann 1966) and cybernetics (Wiener 1961). Our efforts parallel agent-based modeling of physical, social, and socio-economic systems (e.g., Axelrod 1984, 1995, DeAngelis & Gross 1992, Epstein & Axtell 1996, Gaylord & D'Andria 1998, Huberman & Hogg, 1995, Kauffman 1993, Langton 1989).

are adopting.

The task we model explores the mutual effects of events, behaviors, and objects under varying agent properties, event types, group sizes, and task stability. It is a generic *Drosophila* task on which we can begin to tack our theory. The task defined within this model is simple:

> Agents seek specific items in an external search space (e.g., the Internet, a warehouse) and may cooperate (give and receive advice) as to item locations or the quality of other Agents' advice.

What complicates matters is how task and agent properties impact choices, events, affect, and emotions. As this task is communication-based (i.e., all interaction between Agents are communications), then *trust* (*in* other Agents' advice) is a natural cognitive construct to explore within the setting. In our model, trust has both cognitive and emotion-based components. As this task is also about communicating collectives, then *cooperation* (giving advice about *locations*) and *rumor* (giving advice about other Agents' *advice*) are additional organizational constructs to explore within the setting. In our model, generating advice and rumor are behavioral responses that also have both cognitive and emotional components.

In this paper we first describe the three theoretical contributions to the paper. We then present the details of the model and how emotion, trust, cooperation, and deception fit together. We conclude with a report on the results of an initial set of simulation experiments that begin to exercise some components of the model. In these initial experiments we explore how differential trust models, presence of deceptive agents, and rumor impact group performance under stable and unstable task conditions.

An Elicitation Theory

Ortony et al. (1988) present a detailed description of what could be described as a cognitive-based theory of emotional elicitation. We refer to their theory as an elicitation theory (ET) where they posit the overall macro cognitive structure (at an abstract level) linking eliciting conditions (event perceptions) to emotional types and the factors that can influence their intensity.[3]

Eliciting conditions describe the conditions (defined as valenced reactions to situations) under which an emotional type can be triggered. In ET, there are three basic situations that could be perceived (sometimes different from reality) as cueing such reactions: events, agents responsible for events, and objects (*qua* objects).[4]

Each of the three has a primary local variable as a type specification that captures the general form of the relevant valenced reaction. <u>Event-based</u> emotions are triggered and valenced by the *desirability* (pleased, displeased) of the event's consequences relative to one's own *goals* (or the goals of others) without regard to Agent responsibility. <u>Attribution-based</u> emotions are triggered and valenced by the *praiseworthiness* (approve, disapprove) of the actions of the Agent responsible (or believed responsible) for an event with respect to a set of expected *standards* of behavior. <u>Object-based</u> emotions are triggered and valenced by the *appealingness* (appealing, unappealing) of the properties (observed or imputed) of the object involved based on predispositional *attitudes* toward the object (or object type) without regard to Agent involvement, standards, or goals.[5]

Additional variables (global and local) can have effects on the intensity of the experienced emotional type (comprising an inheritance hierarchy), and certain emotional types can be conjuncted under specific conditions (Attribution with Event-based). Our model addresses a subset of the theoretical structure posited by ET as our Agents react to the actions of Agents and consequences of certain Events (not to objects). How the relevant components of ET fit together and are represented computationally will be discussed in the Method section.

An Adaptation and Response Theory

In our model Agents experience events and, under certain conditions, emotional states are invoked (per ET). However, our model requires initial valued states (that permit analysis of threshold levels), adaptation of values with respect to experienced events (e.g., affective and cognitive impact of Agent messages), behavioral choices (e.g., to give or accept advice), and subsequent event experiences. Thus it is essential that our model address, for each event, the effects of the event on: (a) the underlying affect structure of the participating Agent, and (b) the behavioral choices of the Agent. This is handled by Affect Control Theory (Heise 1992, 1987, 1979).

---

[3] In general, we often interact with events around us quite affectively (i.e., in general "positive" or "negative" type reactions). This theory makes the argument "…whether or not affective reactions are experienced as emotions depends upon how intense they are" (p. 20). Thus, implicit in the eliciting condition is the notion of *sufficient* intensity in the context experienced.

[4] Though emotional onsets require varying degrees of cognition, it is important to note that this does not always require deliberations that are conscious to the Agent. Rather, "…they (emotions) are determined by the structure, content, and organization of knowledge representations and the processes that operate on them" (p. 4).

[5] Note that both events (e.g., an influenza injection) and agents (e.g., a newborn baby) can be (situationally) viewed as objects (eliciting dislike and like, respectively). Similarly, objects (e.g., a faulty lawnmower) can situationally be viewed as agents (e.g., source of bad events). The impact of the differential interpretations is that effects the set of emotional responses available.

It has been shown that three factors underlie affective reactions to social and nonsocial event stimuli (and descriptions), and this structure is remarkably stable across cultural groups (Heise 1987). These (valenced and scaled) factors are: Evaluation (a judgement of "goodness – badness"), Potency (a judgement of "power – powerless"), and Activity (a judgement of "vivacity – inactivity").[6]

Let Events experienced by Agents (and descriptions of Events via communication among Agents) be stipulated in ABO constituent form:

Actor (Agent) – Behavior – Object (of the Behavior)

For example, "The *agent* (A) *deceived* (B) the *coworker* (O)."

Each of the constituent elements of such an ABO event description has an associated 3-tuple set of EPA values (Evaluation, Potency, Activity) associated with it (from the perspective of the recipient of the message). Therefore, any given ABO event description has a 3x3 matrix of invoked values (whatever they might be). These are described as *out-of-context* EPA values and reflect the perceiving Agent's fundamental sentiments toward the elements in isolation, but perhaps defined within a general contextual frame (Carley 1986).

Event descriptions, however, describe agents and behaviors *in-context* and it is such events that alter affect, and consequently emotion and behavior. Affect Control Theory (Heise 1987) "shows how affective meanings of social identities and behaviors are maintained while they control interpersonal perception and social action" (p. 1).

Specifically, it

- defines precisely the underlying scales representing the dimensional values, and does so in a manner that the dimensional scales are operationally meaningful – operators can incorporate multiple scales (quantification);
- specifies how the EPA values associated with the constituent elements (ABO) change in context (adaptation); and,
- describes how subsequent behavior choices are influenced (response).

[6] What this means is that Events, Objects and Agents elicit judgements simultaneously on the three dimensions and, therefore, can vary accordingly. The dimension names (e.g., "goodness – badness") are rough approximations to the underlying constructs, but are sufficiently illustrative (see Heise 1987 for details).

Adaptation.

Empirical studies (Heise 1979; Heise & Smith-Lovin 1981; Smith-Lovin 1987) have demonstrated that it is possible to generate a set of adaptation equations that define how event descriptions alter the EPA values of the constituent (ABO) elements.[7]

Consider a simplified example adapted from Heise (1979) of how prior Evaluation scores of word/concepts are adapted in the context of an Event description (ABO). Let A_e, B_e, and O_e be the out-of-context initial Evaluation response ratings (called *fundamental*) for the constituent elements (existing prior to the ABO message). Now, the ABO message describing the event has the effect of updating (adapting) the Evaluation scores (called *transient*) for each of the constituent elements in the following manner. Let R_e represent the resulting Evaluation score of the ABO combined message:

$$R_e = k + \alpha A_e + \beta B_e + \delta B_e O_e$$

This result is then factored as a contextual influence in the adaptation equations as:

$$A_e' = R_e + eA_e$$
$$A_e' = R_e + eA_e$$
$$A_e' = R_e + eA_e$$

All coefficients are empirically determined.

Research in the determination of both the adaptation equations and the coefficient ranges has revealed a coherent set of equations (Smith-Lovin 1987, Tables I, II, III), but also two complicating conclusions (Smith-Lovin 1987). First, and as hinted in the above equation set, there are multiplicative interactions between the constituent ABO elements within dimension (i.e., within E, P, and A adaptation equations). In the above example, the B_eO_e interaction accounts for classically reported Evaluation balance effects (e.g., Heider 1946), but other interactions also fall out to explain other effects.

Second, inter-dimensional interactions are defined. That is, the value on one EPA dimension can impact the updated value on another. For example, a B_pO_e interaction (p = potency) in adapting the Evaluation score of an Agent A_e predicts that Agents are evaluated more positively if they direct powerful events (B_p) at powerful objects/agents (O_p) or less powerful events toward less powerful objects/agents.

For the adaptation equations to be useful, it is necessary to either incorporate or craft a set of constituent (ABO) elements for which their EPA scores have been determined.[8]

[7] In addition, the scaling has been constructed so that each EPA dimension ranges from –4.5 to +4.5 (Heise, 1987).

[8] An example of such a set is given in Heise (1979, Appendix).

Response.

Adaptation equations determine how the elements as configured or described in ABO form (i.e., in the context of the event or message) collectively impact the EPA values of the constituent elements. These updated EPA values are the transient values resulting from that experienced event or event description.

As both the fundamental and transient values are on the same quantitative scales, then it is meaningful to define the squared differences between the EPA values of the ABO constituent elements:

$$
\begin{array}{lll}
(A_e - A_e')^2 & (B_e - B_e')^2 & (O_e - O_e')^2 \\
(A_p - A_p')^2 & (B_p - B_p')^2 & (O_p - O_p')^2 \\
(A_a - A_a')^2 & (B_a - B_a')^2 & (O_a - O_a')^2
\end{array}
$$

These values are called *deflections*, and measure the extent to which a specific event deflects transient sentiment from original values – higher values represent more deflection. Note that there are two ways to interpret deflections. On one hand, once an event/message has occurred, this determines the specific deflections *caused* by the event/message.

On the other hand, given a presumed event (not yet occurred), the adaptation equations can be incorporated to *predict* the deflection of an ABO event/message. This is the important point:

> Affect Control Theory argues that choices of subsequent ABO events (to the extent they are available) are those that *minimize the deflection from the fundamental sentiments*. In effect, this is an effort to select events that confirm the fundamental sentiments.[9]

Therefore, affective judgements from described or experienced events can be mapped into subsequent behavioral choices, and behaviors (as well as affective adaptations) can be predicted and modeled.

Thus Affect Control Theory provides explicit theoretical and computational mechanisms to augment the cognitive elements of ET, and afford the ability to link events (E) to behaviors (B). When sufficient threshold values as determined by the adaptation equations are exceeded and the context defined in the cognitive structure are satisfied (with respect to the ET model of elicitation), then a valenced response to the situation is activated. We refer to this as the EB Agent model (see Figure 2).

As we can see, there are sufficient theoretical apparati to address the links between observed events and subsequent emotionally influenced Agent behavior. What happens in the context of multiple Agents interacting over time? That is, how can we link Agent behavior to group "emotion" as well as group "behavior"? And, why? Toward that end, we are attempting to address the situation depicted in Figure 3.

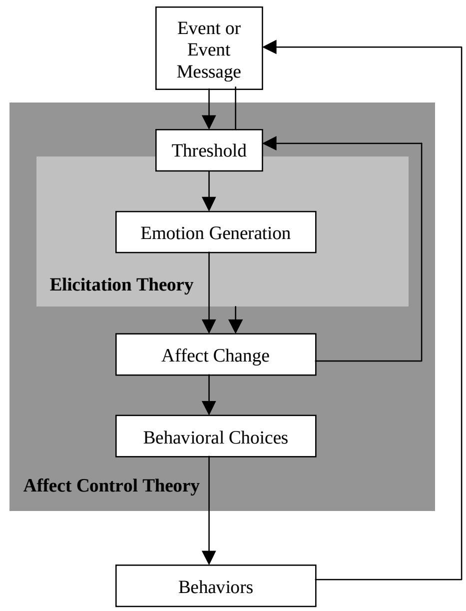

Figure 2. The EB Agent Model

Cooperation, Rumor, and Trust

As we have noted, our models are characterizations of a multi-agent search task. Agent cooperation in these models is defined solely in terms of communication. The impact of emotion influences the choice of cooperation (i.e., communication) behaviors. As such, we explore the constructs of cooperation, trust, and rumor.

Note that it may be important to determine whether the reference subject group (of the Heise data) is sufficiently similar to the target group (of current application).

[9] An corollary to the theory, and seems to be supported by empirical work, is that events that cause little deflection are seen as more likely to occur than event that cause large deflections (Heise 1979; Heise & Mackinnon 1987).

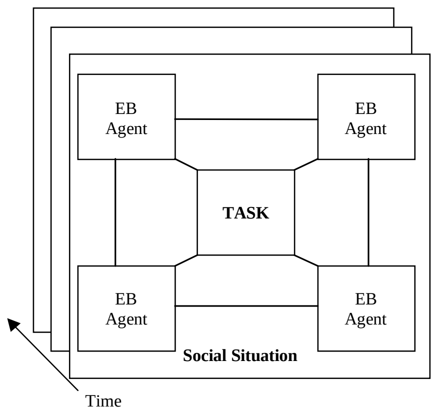

Figure 3. The Multi-Agent EB Model

Cooperation.

In our model, agents can seek (and provide) cooperation in the form of providing advice of two types. First, agents can provide information to facilitate the search goal of another agent by communicating some of its experiential knowledge – whether the agent has seen the item (this requires a task memory for the agent). Second, an agent can provide information that characterizes the agent's experiences with another agent – whether the agent in question has provided helpful location advice (this requires a social memory for the agent). In general, advice is always in response to a question from another agent. An agent chooses to request advice as a search strategy, but an agent chooses to provide advice based on emotionally influenced issues regarding agent behaviors.

Because of the critical role of cooperation to an agent and an agent group, advice becomes an important individual and social construct that serve individual as well as social goals. Consequently, we include a *deception* construct for advice. Thus, agents can not only cooperate or not (i.e., give advice or not), but can also "anti-cooperate" (i.e., interfere) for either individual or social reasons. Accordingly, trust in advice must be considered.

Trust.

If cooperation in the form of advice is a critical component of behavior, then it is important and functional for agents to consider whether or not the source of advice is believable. In our model, this simply means that if Agent *i trusts* Agent *j*, then Agent *i* will *believe* any advice from Agent *j*.

Trust has a host of definitions. In our work we generically define (interpersonal) trust as the Agent's ability, given attributional information, to act on predictions and make predictions that other Agents will act in a cooperative manner. Trust is primarily a cognitive construct (based on deliberation) and secondarily an emotional one (Prietula & Carley, in press). The cognitive components of trust is based on default assumptions (of other agents), and direct experience (with agent communications). For example, we might assume that, at least initially, all agents trust each other. If an agent does not provide correct advice, then that agent might not be trusted or might be suspect (depending on the trust-algorithm for the agent).

When humans interact with agents, the picture becomes more complex. In particular, when advice is characterized from a computer agent, humans react differently in their construction of trust judgements (Lerch, Prietula & Kulik, 1997). This appears to be based on a quickly emergent series of attributional effects centered around the attribution of knowledge to the system, and can even impact a construct of "machine faith" (Lerch, Prietula & Kim, 1999).[10]

Emotion impacts trust when affect responses exceed thresholds under certain circumstances. Extending the above example, if an agent does not provide correct advice, then an emotional response could be activated if either the importance/significance of the event is large, or if the expectations of the agent are radically violated. On the other hand, an assumption of trust might be made as an emotional response to advice from a trusted source that performed a significantly approvable action (e.g., good advice).[11]

Emotion and trust are highly contextual and may have both individual and social consequences. The contextuality is (properly) suggested by the definition provided by Ortony et al. (1988). Thus, the exact nature of how emotion and trust unfold depends on the task, agents, and social situation.

Rumor.

We include rumor in this model as both a behavioral choice and as an organization construct as it reflects a very real social mechanism of communication. Current communication technologies alter the speed and scope of their distribution and (presumably) their organizational effect. Once whispered or dispensed on the telephone, rumors now are broadcast real-time on firm's email or the nation-world Internet in seconds.

In their classic study of wartime rumors, Allport and Post-

[10] In Lerch, Prietula and Kulik (1997), a series of human experiments revealed how trust in a computer-based agent could be defined and influenced. Lerch, Prietula and Kim (1999) is a follow-up series of experiments where they develop an instrument to assess and predict trust responses in a computer-based agent.

[11] Cognition and emotion work together in effecting the three components of trust (Remple, Holmes & Zanna 1985): predictability, dependability, and faith. Predictability refers to the most concrete and observable dimension of trust – predicable agent behavior over time in a presumably *stable* environment. Therefore, informing an agent on the stability or instability of the environment could impact these attributions.

man (1965) provide a definition of rumor as "…a specific (or topical) proposition for belief, passed along from person to person, usually by word of mouth, without secure standards of evidence being present" (p. ix). They argued that there are two basic conditions for rumor: (1) the content involves something of "importance" to the speaker and listener, and (2) the truth is ambiguous.[12]

Rosnow and Fine (1976) offer a succinct definition of rumor as "a proposition that is unverified and in general circulation." Thus, the truth or falsity of a rumor is not the issue, for truth or falsity is unknown; rather, it is that truth or falsity is not immediately verifiable and that the proposition be dispersed. Rumors differ from other sorts of social story exchanges (e.g., legends) in that they address current events, are about specific facts with respect to those events, and are intended to be considered for belief (Kapferer 1990).[13]

Our interpretation of rumor, in fact, overlaps with what is called "gossip." Rosnow and Fine (1976) make a distinction between rumor and gossip in that *gossip* involves "less important," more personal issues, that can be either true or false, with a shorter communicative life-cycle, that serves more ego-based and social-exchange functions. *Rumor*, on the other hand, involves some sort of closure-seeking, is not verified, often concerns more "significant matters," and even serve to function as a mechanism for group problem solving (Shibutani 1966). We see both types occurring in a work setting. A rumor can exist regarding possible lay-offs (facts are unknown, but the matter is significant) while gossip can surround the discussion (e.g., Who is spreading the rumor? How reliable are they?).

In our model, rumors are communications of two types: information about a location (similar to the definitions of rumor above), but also information about other agents (similar to the definition of gossip above). Our interpretation of rumor subsumes that of task-related "institutional gossip" as we see this as an important component of defining and maintaining *informational coalitions*.[14]

Description of Model

The metaphor for this task is a group of agents searching a warehouse for particular items. We will explain the individual elements of the model as Tasks, Agents, and Social Situation.

Task.

The task involved a set of agents, where each agent sequentially acquires an order for an item at a particular location (the order stack), then searchers for that item in a series of other locations (in the warehouse), retrieves that item when found, places the item at another specific location (conveyor), and then proceeds back to the order stack for the next request. The general form of this is shown in Figure 4.

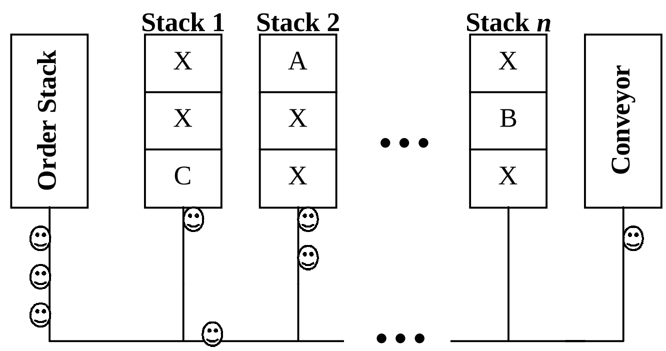

Figure 4. Structure of Warehouse Task

An agent can determine if an item is in a location (stack) only by going to that particular stack's queue. If the queue

[12] In fact, they offer a simple formula for predicting the amount or intensity of rumor in circulation as: $R \sim i \times a$, where i is the importance of the subject and a is the level of ambiguity of evidence on the subject. Models of dissemination of rumors as, for example, birth processes through a population are well documented (e.g., Lave & March 1993), though not quite relevant for this simulation (i.e., spread is not a parameter of the model in this paper).

[13] Although modern "urban legends" are told to be believed (Degh & Vazsonyi 1976), they differ in that they are often derivatives of earlier legend (Bennett 1985) and serve more general entertainment goals (or social goals attainable through entertainment) than task-specific goals as are addressed in our model.

[14] This interpretation comes from three sources. First, work by Dunbar (1996), who argues that gossip is actually a linguistic derivative of grooming behaviors that serve an important social function of maintaining of social bonds, as well as communicating information about the behavior of others (e.g., free riders). In work settings, similar bonds are made with the establishment of organizational coalitions. Consequently, institutional gossip (i.e., gossip whose subject content is related to the agent in the workplace) is a plausible form of coalition maintenance (mutual grooming by exchanging information to facilitate task performance) and control (communicating information about unreliable coalition members who inhibit task performance). Second, Simon (1990, 1993) promotes an argument that in the context of bounded rationality and a requirement for social groups, evolution has promoted a trait of *docility* which is "the tendency to depend on suggestions, recommendations, persuasion, and information obtained through social channels as a major basis for choice." The side effect is that a mechanism exists that can benefit the fitness of the individual (i.e., information about experience contributes to fitness) and the group (i.e., bounded rationality inhibits unambiguous determination about the quality of the information, so the group can occasionally extract altruistic-like behaviors). Thus, group level benefits can be realized by exploiting bounded rationality of individuals. Third, this seems to be supportable by recent research on multi-level selection theory, where group-level adaptive traits in general, and altruism in particular (Sober & Wilson 1998, Wilson & Sober 1994)

is empty (i.e., no other agents are there), then the agent can access the stack and retrieve the item. If the queue is not empty, the agent must wait until it is before the stack can be accessed.

Parameters of this simple task can be altered to systematically modify critical properties, for example:

- number of items and stacks (size of search space);
- replicated items in stacks (size of search space);
- weight of items (nature of search space – single vs. multiple agents required, or items valued by weight);
- volatility of item locations (stability of search space);
- replicated items in order stack (volatility of search space);
- differential value of items (nature of search space).

Agents.

As noted, our agents are boundedly rational, goal-driven agents (though variations will occur as manipulations). In addition, the agents will have the capacities (as knowledge) to address the requirements of various social situations defined in Figure 1 (again, as manipulations). Agents are *maximally defined* with the following properties:[15]

- *item memory* (can recall items in stacks that the agent has visited);
- *event memory* (can recall non-social events);
- *social memory* (can recall social interchanges with other agents);
- *task goals* (define the task objective or objectives for the agent);
- *social goals* (define the constraint or constraints on social behavior for the agent within the task setting);
- *individual goals* (define the objective or objectives of the individual agent as the task unfolds);
- *behavioral typology* (defines specific advice, trust and rumor propensities);
- *communication capability* (can ask and answer questions regarding item locations and agents);
- *expectations and attributions* (default and initial state beliefs, assumptions and attribution values regarding the agent, the task and other agents);
- *emotional typology* (propensity to emotional reactions under situations defined by threshold values);
- *emotional state* (current emotional status in terms of ET);
- *behavioral states* (the behavior repertoire options of the agent in terms of fundamental and transient EPA values);
- *processing capacities* (defining the general constraints on internal deliberation efforts, such as memory limits); and,
- *effort map* (defines how much effort, in terms of processing capacity, is in use and how).

[15] Different social situation manipulations will use different subsets of these constructs. Furthermore, some of these constructs have subconstructs. For example, the relevant set of global and local variables required for ET (Ortony et al. 1988).

In addition to these agent properties, agent messages also have properties (e.g., source of message) depending on the particular message type (e.g., rumor, question, advice).

Social Situations.

The set of social situations represent a series of manipulations that are based on the categories described in Figure 1. As we move through the categories, more knowledge demands (of the social situation) are made on the Agents. In addition to knowledge, each social setting has a slightly different implication for the primary manipulations of interest – advice, rumor, trust and deception. We summarize them as:

- *Nonsocial*. Agent is alone in the task; individual and task goals introduced here; no social knowledge or goals.
- *Nonsocial Multiple Agents*. Multiple agents introduced here; no inter-agent communication; no social knowledge or goals.
- *Real-time Interactive*. Inter-agent communication and rudimentary social and organizational knowledge and learning introduced here; no social goals; advice, rumor, and trust judgements are experience-based.
- *Social Structure*. No social goals; advice, rumor and trust judgements are experience-based and social-structural-expectation based.
- *Social Goals*. Social goals introduced here; advice, rumor and trust-judgements are experience-based and social-goal based.
- *Cultural-Historical*. Historical- and cultural-based norms, beliefs and values (and sanctions for violations) introduced here; advice, rumor, and trust-judgements are experience-, social goal-, and historically-culturally-based.

Interactions.

Interactions occurring among the Task, Agent and Social Situation are important to defining meaningful properties and manipulations, as well as consideration of the utility and etiology of those properties, as shown in Figure 5. Articulating the three elements is not simply a way of viewing the situation, but rather a necessary and sufficient way of viewing organizational and group behaviors (Carley & Prietula 1994). To even define group and organizational "settings," all three elements must be addressed, or explanatory power is decreased (because of the increased degrees of freedom for the unspecified constructs).[16]

Consider the number of agents in the task. Larger numbers

[16] Versions of such interactions as constraints on theory and explanation have been well-argued elsewhere, ranging from psychology (e.g., Hoffman, Feltovich & Ford 1997, Jenkins 1979, Newell 1973) to economics (e.g., Schelling 1978).

of agents will (presumably) have differential effects depending on the type of task and the social situation. Agents in a Nonsocial (Multiple Agent) situation view these additional agents simply as additional objects to negotiate as they pursue their own goals. As they do no communicate, there is no ability to exchange information and share knowledge. Agents in a Real-time Interaction situation have the ability to share information and possibly reduce agent effort. Agents in a Social Structure situation may be impacted by where (in the social structure) the new agents reside. Agents in a Social Goals situation are affected in terms of the social goals that the new agents bring to the task. Finally, agents in Cultural-Historical Situation bring norms, values and beliefs that may or may not be consistent with, and therefore helpful to, the existing set of agents.

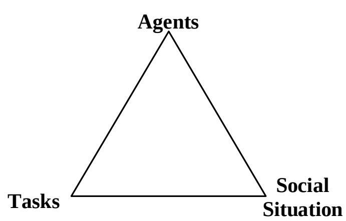

Figure 5. Required Elements of Group Behaviors

A Simulation Experiment

We describe a simple simulation experiment where we examine the behavior of boundedly rational emotional agents under varying task conditions in specific social settings. [17]

Task.

The warehouse task was implemented with ten locations and thirty unique items that were systematically assigned in decreasing modulo fashion starting at Stack 1 (with Item30, Item20, and Item10) through Stack 19 (with Item21, Item11, and Item1). This assignment was constant for all simulations. One critical task property was manipulated, environmental stability, as follows. If a task environment was *stable*, agents could access items in stacks without disrupting other items in the stack. If a task environment was *turbulent*, agents accessing a stack with multiple items would cause disruption by random placement of non-retrieved items to other stacks.

Social Situation.

The social situation was a Real-Time Interactive setting where agents could communicate with other agents. No particular social structure is defined, so judgements are based on direct or communicated experience. Rumor structures are supported as a communication structure, but the lack of social structure imposes no specific group constraints or effect. Task and individual goals exist, but no social goals exist.

Agents.

Five agents were used in all simulations, and all simulations began with the agents queued at the OrderStack. For all simulations an agent could take only one order at a time. Three (indeed interrelated) manipulations were made to the agent models.

- Agent *cooperation*,
- Emotionally-influenced *trust* models, and
- Propensity to assert and listen to *rumor*.

Cooperation. As noted, cooperation in this task is defined as provision of information (advice) in response to a request from another agent. Cooperation can be truthful, or can involve deception (i.e., lies). Recall that there are no social goals in the Real-Time Interaction setting, and task and individual goals dominate choices and interpretations. This is significant, in that addressing self-interest goals appear to the dominant factor in deception choices (Grover 1993, 1997). In a social setting dominated by interaction, information and self-interest, deception is not unlikely.

Deception in our model can be interpreted as a betrayal of expectations, as all agents in all conditions have a default presumption of honest cooperation from other agents. Elangovan and Shapiro (1998) make distinctions among *accidental* betrayal (betrayal event, but absence of intent), *intentional* betrayal (betrayal event, intent present), and *opportunistic* betrayal (betrayal event, intent present but arises situationally and is justified in context). In our model accidental betrayal can arise when honest agents provide advice in turbulent tasks. Intentional betrayal occurs when the default cooperation choice of an agent is deception. Opportunistic betrayal occurs as an emotional choice to change from an honest to a deceptive agent under certain trust events (described next).

In this simulation two types cooperative agents were defined: honest or deceptive. Agents remain that way throughout the simulation. We also vary the number of deceptive agents in the group (for the other manipulations), ranging from zero (no deceptive agents) to five (all agents are deceptive).

Emotion and Trust. In this simulation, the ET components were not included; rather, four specific response types were defined in terms of their reaction to negative events (advice from agents). The set of response types selected subsume differences in importance of events to evoke emotional responses, as well as differences in behavioral responses to events when they are deemed important (i.e., threshold values exceeded). The focus of this aspect is the *behavioral responses* to emotional states defined in terms of trust judgements (believe or not) in agent advice.

[17] The simulation was written in CLIPS 6.0 (CLIPS 1993).

The trust models used in the simulation are the following:

- Model 1 – Trusting. These agents do not see events as sufficiently important to generate an emotional response to alter trust behavior. The always view other agents as honest and trust them. For these agents, emotional states (and thus responses) are not relevant.
- Model 2 – Forgiving. These agents' exhibit emotional responses that are based on bad advice, or rumor, about other agents. For these agents, it takes two pieces of bad advice in a row from an agent before that agent is judged as deceptive. On the other hand, two pieces of good advice in a row can redeem the agent (and be once again viewed a trustworthy).
- Model 3 – Reactive. These agents are similar to the Reactive agents in that only one piece of bad advice causes them to make judgements of deception. Once judged, however, these agents do not forgive. Emotional reactions to events are quickly translated into enduring dispositional judgements of agents.
- Model 4 – Distrusting. These agents do not trust any other agent.

Not only does a judgement of deception (of another agent) cause an agent to *ignore* advice, but it also causes the agent not to *provide* advice to a deceptive agent that is requesting advice.

Rumor. As noted, rumors can play an important part in a group or organizational setting as surrogate experiences and attributional influences. In this simulation, rumors are incorporated as follows:

- If some Agent *i* received bad advice from some Agent *j*, and Agent *i* is a gossip, then Agent *i* starts a rumor that Agent *j* is a liar.
- Any Agent *k* (i.e., in all above models) that detects a rumor about another Agent *j*, will choose not to provide advice to Agent *j*.

Rumors have an immediate effect. Two levels of this variable are incorporated as a property of the agents: assert rumors, do not assert rumors. Perceived rumors do *not* change an Agent's judgement (about the rumored Agent), but only whether to provide advice or not, or to accept advice, or not. However, the Agent originating the rumor has experiences that can alter trust judgements (based on the particular Trust model).

Deception. Agents that are deceptive always attempt to provide false information about locations when asked by other agents. These agents generate a random location when asked, but do not check whether that location is accurate or not (they, in fact, may not know). Thus a small fraction of the time these agents may provide accurate advice. Deceptive agents do not withhold information (they generate false information), but they are susceptible to receiving false information, so they can generate rumors as do non-Deceptive Agents.

Design.

Manipulations were made at the group level (i.e., homogeneous Agents within manipulation). The manipulations were:

- Task. Two levels of Environmental Stability were examined (stable environment, turbulent environment).
- Agents. Four Trust models (Trusting, Forgiving, Reactive, and Distrusting), Two levels of Rumor (asserts, does not assert), and six levels of Deceptive Agents (0. 1, 2, 3, 4, and 5) in any 5-agent group.

We report on the following sample of dependent variables:

- *Organizational effort* (total number of effort by the group);
- *Information Withheld* (an agent knows the answer to a question from another agent, but does not answer because there is a rumor about the questioning agent, or the agent is judged to be deceptive);
- *Structural Duration* (average duration of positive trust relations expressed as percentage of maximum cycles);
- *Conflict* (relations that have been judged as deceptive as a percentage of the total coalition).

Results and Discussion

Main Effects. If we make the (admittedly broad) assumption that the set of models and task/agent properties we have selected are somewhat representative of a set of random variables taken from a population of models and properties, we can take a high-level view of main effects and interactions across that set. Table 1 presents the main effects obtained.

Table 1. Main Effects

	Effort	**Info Withheld**	**Duration**	**Conflict**
Trust	↑*	↓***	↑***	↓***
Turbulence	↑***	ns	ns	ns
Deception	↑***	↑**	ns	ns
Rumor	ns	ns	↑***	↓***

*p < .05, **p < .01, ***p < .001

Increasing the trust models of the group results in subsequent increases in overall effort and more durable information coalitions, and subsequent decreases in the information withheld and conflict. Increasing the turbulence of the environment has a single effect of increasing effort. As more deceptive agents are added to a group, both effect and information withheld increases. Finally, if rumors are incorporated, the duration of the information coalitions and decreases the conflict of the group.

These models impact all four dependent variables, so we can explore their differential impacts across dependent variables, by a post-hoc analysis (Duncan's multiple range test) of the results (Table 2).

From this analysis, we can see that there does exist variance in the effects of the models, and these effects differ across dependent variables. Table 2 is read as follows. Agents that are Trusting (model 1), score significantly higher in Effort and withhold more information than Agents that are Distrusting (model 4), have more stable information coalitions than all of the other models (2, 3 and 4), and have lower conflict scores that the other models (2, 3 and 4).

Table 2. Model Differences

Differs from…	**Effort**	**Info Withheld**	**Duration**	**Conflict**
1 Trusting	4	4	2,3,4	2,3,4
2 Forgiving	4	4	1,3,4	1,4
3 Reactive	---	4	1,2,4	1,4
4 Distrusting	1,2	1,2,3	1,2,3	1,2,3

Overall, most of the effects are what one may expect, but bearing in mind the nested effects, we can explore further the interactions. However, limits in the values (variance) in some conditions restrict the explorable set of interactions. Consider, for example, the "extreme" behaviors of the Trusting and Distrusting agents – the former is always trustful while the latter is never trustful – so Duration and Conflict values cannot be used. In this analysis, we disregard these extreme models and explore full interactions for the Forgiving and Reactive models only.

Figure 6 and 7 show the interaction graphs for the Forgiving and Reactive models respectively, exploring how Deception, Environmental Stability, and Rumor impact organizational Effort.

Consider the left two panels (with no Deceptive Agents and one Deceptive Agent) of Figure 6. Turbulent (uncertain) environments increase the organizational effort. Rumors have little effect in turbulent environments, but begin to have some effect in stable ones (with one Deceptive Agent). As Deceptive Agents are added to the mix, benefits from rumors are reaped in stable environments and, to a lesser extent, in turbulent ones.

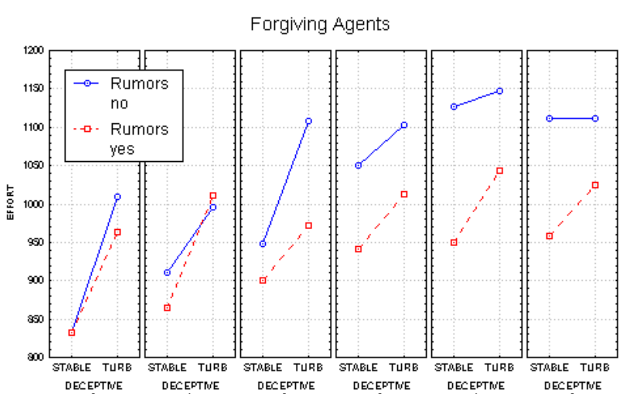

Figure 6. Interactions for Forgiving Agents on Effort

Turning to Reactive Agents (Figure 7), the benefits from rumor are accrued but the benefits are more discontinuous. In addition, the general level of effort is attenuated as the number of Deceptive Agents increase. When the number of Deceptive Agents is three or five, there seems to be a slight Environment by Rumor interaction.

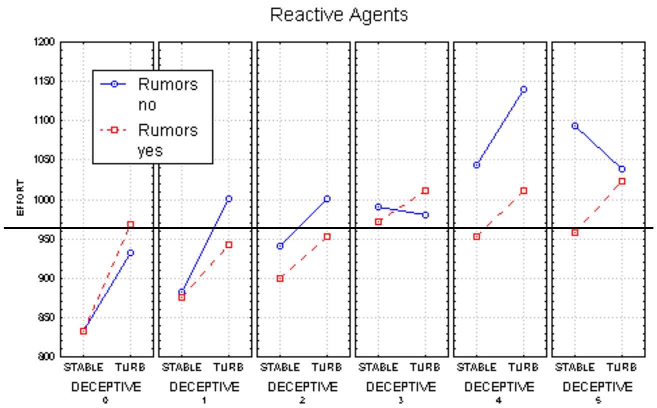

Figure 7. Interactions for Reactive Agents on Effort

Thus, Forgiving Agents benefit more from rumor in Deceptive and Turbulent environments than do Reactive Agents, as Reactive Agents not only actively spread rumor, but are more immediate in their behavioral changes with respect to advice as the likelihood of correct advice decreases (via Turbulence or Deception).

Figures 8 and 9 show the interaction graphs for the Forgiving and Reactive models respectively, exploring how Deception, Environmental Stability, and Rumor impact Information Withheld.

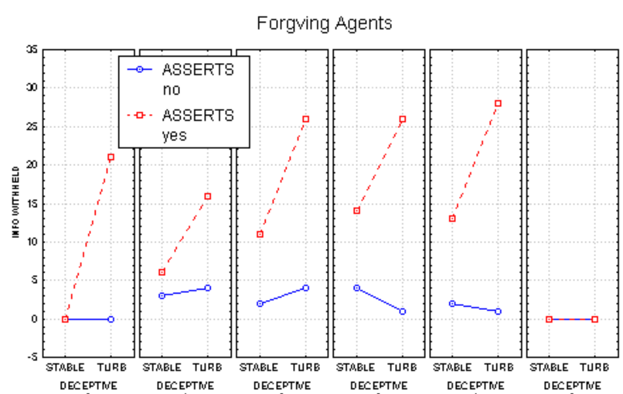

Figure 8. Interactions for Forgiving Agents on

Information Withheld

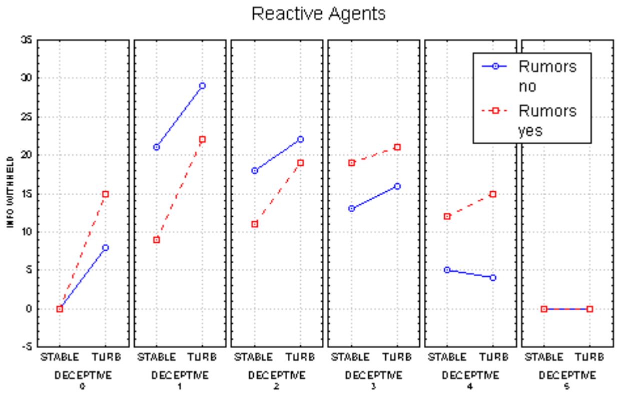

Figure 9. Interactions for Reactive Agents on Information Withheld

Note that information withheld means that a given Agent *i* perceives a request for advice from another Agent *k* and Agent *i* knows (or presumably knows) information of value to the questioning Agent *k*. If there is a Rumor about Agent *k*, the advice is withheld by Agent *i*. If Agent *i* has previously made a judgement about Agent *k* as being untrustworthy, the advice is again withheld by Agent *i*. On the other hand, recall that Deceptive Agents do *not* withhold information (though they are susceptible to bad advice, make judgements of other agents, and generate rumors). Thus, Information Withheld is a direct function of rumor about a given Agent for non-Deceptive Agents.

First, we can examine the behavior without Rumors. In Figure 8, the stability of the environment has little effect on Information Withheld. Without rumors, Forgiving Agents must rely on direct experience, and that experience includes higher tolerance for bad advice, thus a general reduction in Information Withheld. Additionally, as the number of Deceptive Agents grows, the number of Agents withholding Information decreases (recall Deceptive Agents do not withhold Information).

The graph for the Reactive Agents (Figure 9) is somewhat more involved, but the same phenomena hold. Like the Forgiving Agents, Reactive Agents respond to advice and when that advice is bad, they alter their trust judgements. If such Agents are judged as untrustworthy, then they will not receive responses to their requests for advice (except by Deceptive Agents). Unlike Forgiving Agents, Reactive Agents are less forgiving and judge more Agents as untrustworthy; therefore, they withhold information when asked. Without rumors, these Agents rely on experience and their experience generates judgements and withholding of information. However, as the number of Deceptive Agents grows, the number of Agents withholding Information decreases, causing the overall decreasing level.

In order to examine the cases with Rumor, it is necessary to recall that information is withheld (by non-Deceptive Agents) because (1) the questioning Agent has been judged untrustworthy (via direct individual experience), or (2) there is a rumor that the questioning Agent is untrustworthy, where condition (1) dominates condition (2). Furthermore, as rumors inhibit direct experience, *rumors inhibit judgement changes*. For Forgiving Agents, their judgements are generally positive (Figure 8), so condition (1) above is generally not met. However, when Rumor is introduced, condition (2) above is met and persists as the number of Deceptive Agents increases. In fact, as the number of Deceptive Agents increases, the decreasing number of non-Deceptive Agents accounts for Information Withheld.

For Reactive Agents, low numbers of Deceptive Agents result in judgement changes that dominate the effects of Rumor. As direct experience causes more negative judgements, less Agents are likely to listen to advice, experience bad events, and spread rumor. Recall that Rumors are derivative of following bad advice, so when less Agents follow less advice, less rumors are spread and, therefore, less rumors are taken (see Figure 10).

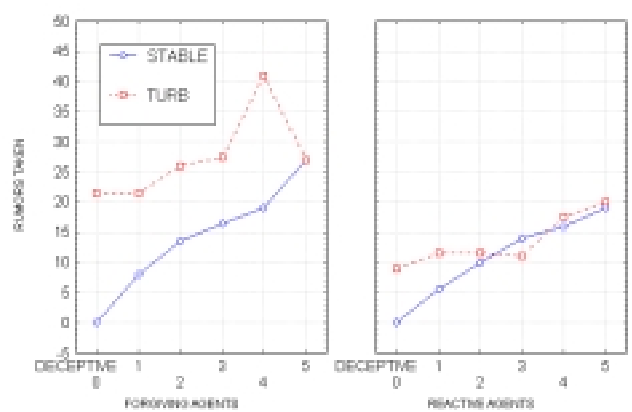

Figure 10. Rumors Taken by Agents

Figures 11 and 12 show the interaction graphs for the Forgiving and Reactive models respectively, exploring how Deception, Environmental Stability, and Rumor impact the average Duration (as a fraction of maximum simulation cycles) of the information coalition members. For example, in Stable environments with no Deceptive Agents, the Information Coalition is maximally intact at 1.0.

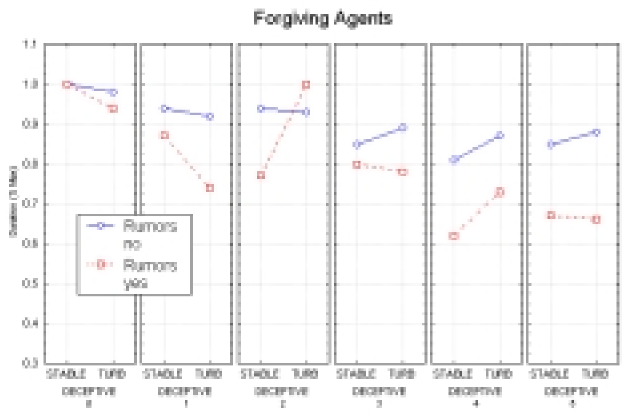

Figure 11. Interactions for Forgiving Agents on Duration of Information Coalitions

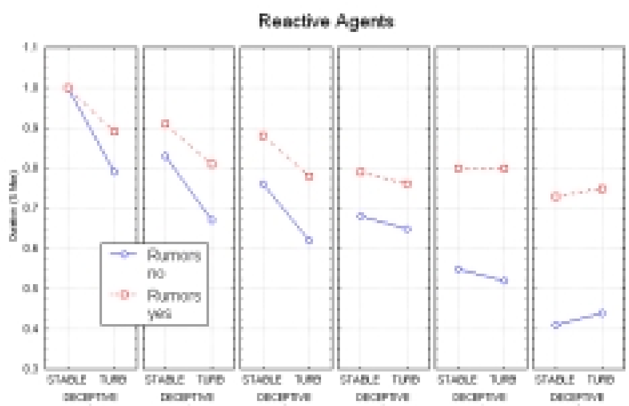

Figure 12. Interactions for Reactive Agents on Duration of Information Coalitions

In Figure 11, it can be seen that as the number of Deceptive Agents increases, the general duration of an information coalition of Forgiving Agents does not decline appreciably, and Turbulence does not seem to be a factor. Rumors have a general effect of reducing the Duration situationally. On the other hand, Reactive Agents (Figure 12) that do not use rumor have a general reduction in the Duration of the coalitions as the number of Deceptive Agents increases. For Reactive Agents, rumor sustains the Duration over increasing Deceptive Agents and turbulent conditions.

Thus, rumors have almost the opposite effects for the different Agent trust types. The reason for this is, again, the differential roles and effects of experience and rumor. Agent trust models are based on experience; that is, trust adjustments are based solely on advice from other Agents. Rumors, as noted, are weaker experience surrogates that inhibit the generation of advice (as cooperation), but attenuate the alteration of judgements. The Forgiving Agents of Figure 11, under Rumor conditions, do not follow advice (because of rumor) and thus cannot forgive the Agents. Rumor disrupts the coalitions. The Reactive Agents of Figure 12, on the other hand, are inhibited from imposing their negative assessments and rumor serves to preserve coalitions.

Finally, Figures 13 and 14 show the interaction graphs for the Forgiving and Reactive models respectively, exploring how Deception, Environmental Stability, and Rumor impact the Conflict (fraction of the information coalition disrupted).

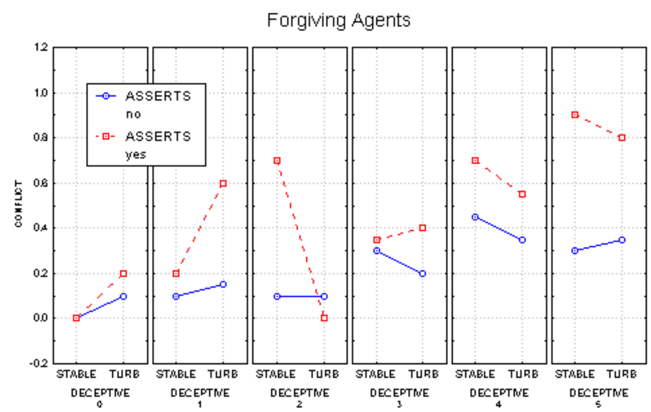

Figure 13. Interactions for Forgiving Agents on Conflict

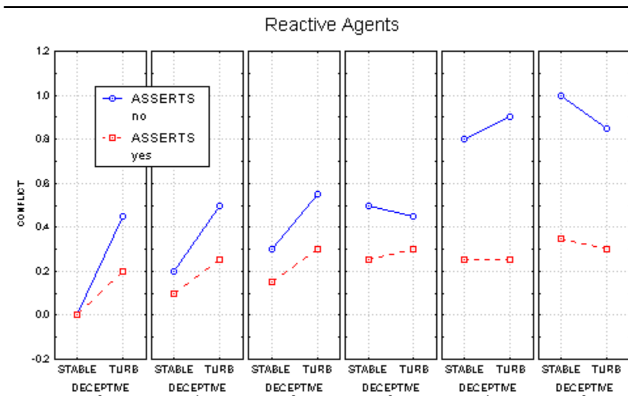

Figure 14. Interactions for Reactive Agents on Conflict

Conflict is closely related to the Duration measure, so the explanations are equivalent. Left to their own calibration devices, Forgiving Agents (Figure 13) will augment their judgements through interaction with the Agents, but this is prevented when rumor structures are present. Similarly, Reactive Agents (Figure 14) are quick to break up coalitions and increase Conflict. Again, rumor inhibits the direct experience and, by our definition, reduces conflict in the coalition.

How this works can be explained by examining the interaction between trust judgements adjustments and rumor. When Agents experience good or bad advice, they alter their trust judgements accordingly, though slightly differently as their trust models differ. Across Deceptive conditions, the general amount of good advice an Agent receives understandably decreases: more Deceptive Agents are generating bad advice and fewer Agents are generating good advice. Rumor, however, does not impact good advice as much as bad, as rumors are spread based on bad advice, not good (reflecting the general nature of rumor). Therefore, the effect is that rumor attenuates bad judgements, which necessarily depend on experience, and this impacts the Reactive Agents' ability to alter trust judgements down within Rumor conditions (Figure 15). This effect is also found in Turbulent environments, though there is an overall effect of less judgements up.

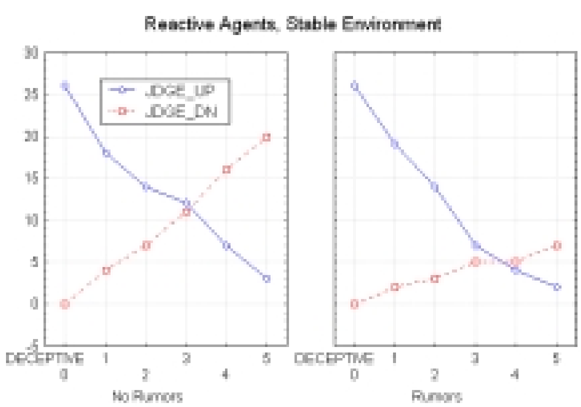

Figure 15. Reactive Agents and Judgements

Conclusion

The purpose of this paper was to craft an initial process-level simulation in order to explore how different emotional-trust models of individual boundedly rational and emotional agents would impact group performance that depended, in part, on cooperation in the form of information. We defined four Trust models and explored how their decisions unfolded under a variety of conditions. The general nature of the Trust models were influenced by empirical studies on how humans trust information from a series of different sources: peers, experts, and intelligent (computer-based) agents. We also described how two models of emotion could be woven together to generate emotion-based responses, though we did not implement this in this study.

Does trust matter? Overall findings suggest that different trust models have different organizational effects (recall Tables 1 and 2) and we focused our results on two variations, Forgiving and Reactive Agents, and explored how groups would differ under varying conditions of Rumor, number of imposed Deceptive Agents, and whether the task environment was Stable or not. The most interesting results seem to occur with the Rumor construct and its interactions with the other manipulations.

Do rumors matter? Rumors indeed matter, but not the same way to all groups (and we are talking about groups of Agents).

- For Forgiving and Reactive Agents, as the number of Deceptive Agents increases in a group, rumor can reduce their general impact on increasing organizational effort;
- For Forgiving Agents, rumors increase the amount of information withheld (cooperation via advice) as the number of Deceptive Agents increases, but this effect does not carry over to Reactive Agents;
- For Forgiving Agents, rumors reduce the average size of an information coalition, but for Reactive Agents rumor increases the average size of the information coalition;
- For Forgiving Agents, rumors can increase the amount of group conflict, but for Reactive Agents rumor can reduce the group conflict.

How do rumors matter? This, perhaps, is the most revealing result of our simulation. Overall, rumors act by disseminating Agents' negative experiences (i.e., reactions to receiving bad advice) which, in turn, reduces the likelihood of a "deceptive" Agent deceiving other Agents (as rumors induce them to avoid advice from the offending Agent). On the other hand, by not taking the advice of an Agent, direct experiences cannot ensue that may alter the trust judgements of the Agent (up or down). Thus, one can think of adherence to rumors as a group trait, so to speak (i.e., do not supply information to an agent when there exists a rumor), while specific evaluations about an Agent's trustworthiness, based on direct experiences with that Agent, as an individual trait.

From the perspectives of the Agent types, Forgiving Agents have a "maladaptive under-reaction" to perform in extremely uncertain information environments (Turbulence and high Deception), so a group trait (rumor) facilitates the reduction of the information coalitions, and reduces the overall effort of the group (see Figure 6). These Forgiving Agents are prevented for forgiving other Agents (depending on general timing issues) based on their individual preferences for adjusting trust. Rumor facilitates the destruction of the coalition.

Reactive Agents possess individual traits to quickly terminate information coalition membership, but under certain conditions this might be an "maladaptive over-reaction" and rumor serves as a group behavior that mitigates that trait. Acting in the opposite manner, rumor inhibits Reactive Agents from direct experiences that would result in coalition termination. Rumor facilitates the preservation of the coalition. Overall, rumor has a dampening effect on the underlying trust judgements.[18]

This work is embryonic and necessarily constrained. Our general strategy, however, remains the same: determine the empirical foundations of a plausible set of behavioral functions that can account for the collective phenomena under study. The next steps seek to directly incorporate the other specific components (e.g., Affect Control mechanisms, ET structures), scale up (and down) the size of the group, altering the forms of agents, and proceed with human-agent calibration and predictions studies.

Acknowledgements

We would like to thank Richard Burton (Duke), Benoit Morel (Carnegie Mellon), Andrew Ortony (Illinois), and Ray Levitt (Stanford) for their input and suggestions.

References

Albin, P. (1998). *Barriers and Bounds to Rationality*. Princeton, NJ: Princeton University Press.

Allport, G. & L. Postman, (1965). *The Psychology of Rumor.* New York, NY: Russell & Russell.

Axelrod, R. (1984). *The Evolution of Cooperation*. New York, NY: Basic Books.

Axelrod, R. (1985). *The convergence and stability of culturs: Local convergence and global polarization.* Working Paper 95-03-028, Sante Fe Institute, Sante Fe, NM.

Bates, J. (1994). The role of emotion in believable agents. *Communication of the ACM,* 37, 122-125.

Bennett, G. (1985). What's modern about the modern legend. *Fabula,* 26 (3/4), 219-229.

Canamero, D. (1997). Modeling motivations and emotions as a basis for intelligent behavior. In W. Johnson (ed.), *Proceedings of the First International Conference on Autonomous Agents.* New York, NY: ACM Press.

Carley, K. (1986). An approach for relating social structure to cognitive structure. *Journal of Mathematical Sociology*, 12, 108-116.

Carley, K. & A. Newell (1994). The Nature of the Social Agent, *Journal of Mathematical Sociology*, 19(4), 221-262.

Carley, K. & M. Prietula (Eds.) (1994). *Computational Organization Theory*. Hillsdale, NJ: Lawrence Erlbaum.

Carley, K. & M. Prietula (1994b). ACTS Theory: Extending the Model of Bounded Rationality. In K. Carley and M. Prietula (Eds.), *Computational Organization Theory* (pp. 55-87). Hillsdale, NJ: Lawrence Erlbaum.

CLIPS. (1993). *CLIPS Reference Manual, Version 6.0.* Lyundon B. Johnson Space Center, Software Technology Branch, Houston TX.

Cyert, R. and J. March (1963). *A Behavioral Theory of the Firm.* Englewood Cliffs, NJ: Prentice Hall.

DeAngelis, D> & L. Gross (eds.) (1992). *Individual-based Models and Approaches in Ecology*. New York, NY: Chapman & Hall.

Degh, L. & A. Vazsonyi. (1976). Legend and belief. In D. Ben-Amos (Ed.), *Folklore Genres.* Austin, TX: University of Texas Press.

Dunbar, R. (1996). *Grooming, Gossip, and the Evolution of Language*. Cambridge, MA: Harvard University Press.

Elangovan, A. & D. Shapiro. (1998). Betrayal of trust in organizations. *Academy of Management Review*, 23(3), 547-566.

Epstein, J. & R. Axtell. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. Cambridge, MA: MIT Press.

Gaylord, R. & L. D'Andria. (1998). *Simulating Society*. New York, NY: Springer-Verlag/Telos.

Glance, N. & B. Huberman. (1994). The dynamics of social dilemmas. *Scientific American,* March, 76-81.

Heider, F. (1946). Attitudes and cognitive organizations. *Journal of Psychology*, 21, 107-112.

Heise, D. (1992). Heise, D. R. 1992. Affect control theory and impression formation. Vol. I, pp. 12-17. E. Borgatta and M. Borgatta (Eds.), *Encyclopedia of Sociology*. New York: Macmillan.

Heise, D. (1987). Affect control theory: Concepts and model. *Journal of Mathematical Sociology*, 13(1-2), 1-33.

Heise, D. (1979). *Understanding Events: Affect and the*

[18] Turning back to our earlier footnoted arguments (footnote 14), we may consider these findings as a peek at the interaction between the social and individualistic roles of docility and gossip. Intuitively, one could argue that disseminating gossip (about Deceptive Agents) is a valuable social mechanism to attenuate the impact of Deceptive agents (whatever their motives might be). But the value of the mechanism in Turbulent environments can lead to "false positives" – Agents that may convey wrong, but not deceptive, information, thus diminishing the value of the mechanism. Consequently, it may be appropriate to *alter* one's trust strategy in light of new evidence that information regarding Agent's is insufficient for trust judgement, but sufficient for assessing situational reliability. Individual trust models take that into account by lag effects (i.e., as trust builds, variation of reliability has diminishing effects).

Construction of Social Action. Cambridge, England: Cambridge University Press.

Heise, D. & N. MacKinnon. (1987). Event likelihoods. *Journal of Mathematical Sociology*, 13(1-2), 133-151.

Heise, D. & L. Smith-Lovin. (1981). Impressions of goodness, powerfulness, and liveliness from discerned social events. *Social Psychology Quarterly*, 44, 93-106.

Hoffman, R., P. Feltovich & K. Ford. (1997). A general framework for conceiving of expertise and expert systems in context. In P. Feltovich, K. Ford & R. Hoffman (Eds.), *Expertise in Context*. Menlo Park, CA: AAAI Press/MIT Press.

Huberman, B. & T. Hogg. (1995). Communities of practice: Performance and evolution. *Journal of Computational and Mathematical Organization Theory*, 1(1), 73-92.

Hyder, E., M. Prietula & L. Weingart. (to appear). Getting to best: Efficiency and optimality in negotiations. *Cognitive Science*.

Jenkins, J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. Cermak & F. Craik (eds.), *Levels of Processing in Human Memory*. Hillsdale, NJ: Erlbaum.

Kapferer, J-N. (1990). *Rumors: Uses, Interpretations, and Images*. New Brunswick, CT: Transaction Publishers.

Kauffman, S. (1993). *The Origins of Order*. New York, NY: Oxford University Press.

Langton, C. (ed.) (1989). *Artificial Life*. Redwood City, CA: Addison-Wesley.

Lave, L. & J. March. (1993). *An Introduction to Models in the Social Sciences*. New York, NY: University Press of America.

Lerch, J., M. Prietula & C. Kulik. (1997). The Turing Effect: The nature of trust in expert systems advice. In P. Feltovich, K. Ford & R. Hoffman (Eds.), *Expertise in Context*. Menlo Park, CA: AAAI Press/MIT Press.

Lerch, J., M. Prietula & J. Kim. (2000). *Unraveling the Turing Effect: Measuring trust in computer advice*. Unpublished manuscript.

March, J. (1978). Bounded rationality, ambiguity, and the engineering of choice. *The Bell Journal of Economics*, 9(2), 587-608.

Morignot, P. & B. Hayes-Roth. (1996). *Motivated Agents*. Knowledge Systems Laboratory, Report KSL 96-22, Department of Computer Science, Stanford University, Stanford, CA.

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard.

Newell, A. (1973). You can't play 20 questions with nature and win.: Projective comments on the papers of this symposium. In W. Chase (ed.), *Visual Information Processing*. New York, NY: Springer-Verlag.

Ortony, A., G. Clore & A. Collins (1988). *The Cognitive Structure of Emotions*. Cambridge, England: Cambridge University Press.

Picard, R. (1997). *Affective Computing*. Cambridge, MA: MIT Press.

Prietula, M. & K. Carley (in press). Exploring the effects of agent trust and benevolence in a simulated organizational task. *Applied Artificial Intelligence*.

Prietula, M., K. Carley & L. Gasser (Eds.) (1998). *Simulating Organizations: Computational Models of Institutions and Groups*. Cambridge,MA: AAAI/MIT Press.

Remple, J., J. Holmes & M. Zanna (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, 49(1), 95-112.

Rosnow, R. & G. Fine. (1976). *Rumor and Gossip: The Social Psychology of Hearsay*. New York, NY: Elsevier.

Rowe, N. (1989). *Rules and Institutions*. New York, NY: Philip Allan.

Shibutani, T. (1966). *Improvised News: A Sociological Study of Rumor*. Indianapolis, IN: Bobs Merrill.

Simon, H. (1993). Altruism and economics. *Papers and Proceedings of the 105th Annual Meeting of the American Economic Association*, 83(2), 156-161.

Simon, H. (1991). Organizations and markets. *Journal of Economic Perspectives*, 5, 25-44.

Simon, H. (1990). A mechanism for social selection and successful altruism. *Science*, 250, 1665-1668.

Simon, H. (1976a). Administrative Behavior (3rd Edition). New York, NY: Free Press.

Simon, H. (1976b). From substantive to procedural rationality. In S. Latsis (ed.) *Method and Appraisal in Economics*, Cambridge, England: Cambridge University Press.

Simon, H. (1967). Motivational and emotional controls of cognition. Reprinted in H. Simon (ed.), *Models of Thought*. New Haven, CT: Yale University Press, 1979.

Simon, H. (1956). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99-118.

Simon, H. (1955). Rational choice and the structure of the environment. *Psychological Review*, 63, 129-138.

Smith-Lovin, L. (1987). Impression from events. *Journal of Mathematical Sociology*. 13(1-2), 35-70.

Sober, E. & D. Wilson. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.

Williamson, O. (1975). *Markets and Hierarchies*. New York, NY: Free Press.

Wilson, D. & E. Sober. (1994). Reintroducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences*, 17, 585-654.

Wittek, R. & R. Wielers (1998). Gossip in organizations. *Computational and Mathematical Organization Theory*, 4(2), 189-204.