

# Evaluating a General Model of Emotional Appraisal and Coping

Jonathan Gratch  
University of Southern California  
Institute for Creative Technology  
13274 Fiji Way, Marina del Rey, CA 90292  
[gratch@ict.usc.edu](mailto:gratch@ict.usc.edu)

Stacy Marsella  
University of Southern California  
Information Sciences Institute  
4676 Admiralty Way, Marina del Rey, CA 90292  
[marsella@isi.edu](mailto:marsella@isi.edu)

## Introduction

In our research, we have developed a general computational model of human emotion. The model attempts to account for both the factors that give rise to emotions as well as the wide-ranging impact emotions have on cognitive and behavioral responses. Emotions influence our beliefs, our decision-making and how we adapt our behavior to the world around us. While most apparent in moments of great stress, emotions sway even the mundane decisions we face in everyday life [1, 2]. Emotions also infuse our social relationships [3]. Our interactions with each other are a source of many emotions and we have developed a range of behaviors that can communicate emotional information as well as an ability to recognize and be influenced by the emotional arousal of others. By virtue of their central role and wide influence, emotion arguably provides the means to coordinate the diverse mental and physical components required to respond to the world in a coherent fashion [4].

The model of emotion we have developed accounts for a range of such phenomena. This model has been incorporated into human-like agents, called virtual humans. This technology has been used to create a significant application where people can interact with the virtual humans through natural language in high-stress social settings (see Figure this page) [5-8].

Given the broad and subtle influence emotions have over behavior, evaluating the effectiveness of such a general architecture presents some unique challenges. Emotional influences are manifested across a variety of levels and modalities. Emotion is often attributed to others in response to telltale physical signals: facial expressions, body language, and certain acoustic features of speech. But emotion is also conveyed through patterns of thought and coping behaviors such as wishful thinking, resignation, or blame-shifting. Worse, emotions are frequently attributed in the absence of any visible signal (e.g., he is angry but suppressing it) and these attributions can be influenced by the observers own emotional state. Unlike many phenomena studied by cognitive science, emotional

responses are also highly variable, differing widely within and across individuals depending on non-observable factors like goals, beliefs, cultural norms, etc. And unlike work in decision making, there is no accepted normative model of emotional responses or their dynamics that we can use as a gold standard for evaluating techniques.

In the virtual human research community, current evaluations have relied on the concept of “believability” in demonstrating the effectiveness of a technique: A human subject is allowed to interact with a system or see the result of some system trace, and is asked how believable the behaviors appear; it is typically left to the subject to interpret what is meant by the term. One obvious limitation with this approach is that there seems to be no generally agreed definition of what “believability” means, how it relates to other similar concepts such as realism (or example, in a health-intervention application developed by one of the authors, stylized cartoon animation was judged to be highly believable even though it was explicitly designed to be unrealistic along several dimensions [9]).

We attempt to move beyond the concept of believability and instead evaluate more specific functional questions. In this paper, we illustrate this methodology through two studies currently underway at our lab, where each study illustrates a different line of attack. In the first study, we address the question of process dynamics: does the model generate cognitive influences that are consistent with human data on the influences of emotion, specifically with regard to how emotion shapes perceptions and coping strategies, and how emotion and coping unfold over time. In the second, we address the question of behavioral in-



fluence: do external behaviors have the same social influence on a human subject that one person's emotion has on another person, specifically with regard to how emotional displays influence third-party judgments. In other words, (1) does our computational model create the right cognitive dynamics and (2) does it have the right social impact.

### Appraisal Theory (a review)

Motivated by the need to inform the design of symbolic systems, our work is based on cognitive appraisal theory which emphasizes the cognitive and symbolic influences of emotion and the underlying processes that lead to this influence [10] in contrast to models that emphasize lower-level processes such as drives and physiological effects [11]. In particular, our work is informed by Smith and Lazarus' cognitive-motivational-emotive theory.

Appraisal theories argue that emotion arises from two basic processes: appraisal and coping. Appraisal is the process by which a person assesses their relationship with the environment, including not only current conditions, but events that led to this state and future prospects. Appraisal theories argue that cognitive processes inform this perceived relationship (e.g., planning, explanation, perception, memory, linguistic processes) but that appraisal maps aspects of these disparate processes into a common set of *appraisal variables*. These serve as an intermediate description of the person-environment relationship – a common language of sorts – and mediate between stimuli and response. Appraisal variables characterize the significance of events from the individual's perspective. Events do not have significance in of themselves, but only by virtue of their interpretation in the context of an individual's beliefs, desires and intention, and past events.

Coping determines how one responds to the appraised significance of events and people are motivated to respond to events differently depending on how they are appraised [12]. For example, events appraised as undesirable but controllable motivate people to develop and execute plans to reverse these circumstances. On the other hand, events appraised as uncontrollable lead people towards denial or resignation. Psychological theories have characterized the wide range of human coping responses into two classes. *Problem-focused coping* strategies attempt to change the environment. *Emotion-focused coping* strategies [13] are inner-directed strategies for dealing with emotions. Emotion-focused coping alters one's interpretation of circumstances, for example, by discounting a potential threat or abandoning a cherished goal.

The ultimate effect of these strategies is a change in the person's interpretation of their relationship with the environment, which can lead to new (re-) appraisals. Thus, coping, cognition and appraisal are tightly coupled, interacting and unfolding over time [13, 14]: an agent may "feel" distress for an event (appraisal), which motivates the shifting of blame (coping), which leads to anger (re-

appraisal). A key challenge for a computational model is to capture this dynamics.

### A computational Model

EMA is a computational model of emotion processing based on cognitive appraisal theory and described in detail elsewhere [6, 7]. Here we sketch the basic outlines. A central tenant in cognitive appraisal theories in general, and Smith and Lazarus' work in particular, is that appraisal and coping center around a person's *interpretation* of their relationship with the environment. This interpretation is constructed by cognitive processes, summarized by appraisal variables and altered by coping responses. To capture this interpretative process in computational terms, we have found it most natural to build on the causal representations developed for planning techniques and augment them with decision-theoretic planning techniques (e.g., [15]) and with methods that explicitly model commitments to beliefs and intentions [16, 17]. Plan representations provide a concise representation of the causal relationship between events and states, key for assessing the relevance of events to an agent's goals and for assessing causal attributions. Plan representations also lie at the heart of many autonomous agent reasoning techniques (e.g., planning, explanation, natural language processing). Beyond modeling causality, attributions of blame or credit involve reasoning if the causal agent intended or foresaw the consequences of their actions, most naturally represented by explicit representations of beliefs and intentions. As we will see, commitments to beliefs and intentions also play a key role in modeling coping strategies. The appraisal variables of desirability and likelihood find natural analogues in the concepts of utility and probability as characterized by decision-theoretic planning methods.

In EMA, the agent's current interpretation of its "agent-environment relationship" is reified by the output and intermediate results of those reasoning algorithms that relate the agent to its physical and social environment. We use the term *causal interpretation* to refer to this collection of data structures to emphasize the importance of causal reasoning as well as the interpretative (subjective) character of the appraisal process. At any point in time, this configuration of beliefs, desires, plans, and intentions represents the agent's current view of the agent-environment relationship, an interpretation that may subsequently change with further observation or inference. We treat appraisal as a set of feature detectors that map features of the causal interpretation into appraisal variables. For example, an effect that threatens a desired goal would be assessed as a potential undesirable event. Coping sends control signals to auxiliary reasoning modules (i.e., planning, action selection, belief updates, etc.) to overturn or maintain features of the causal interpretation that yield individual appraisals. For example, coping may resign the agent to the threat by abandoning the desired goal. Figure 2 illustrates a reinterpretation of Smith and

Lazarus' cognitive-motivational-emotive system consistent with this view. The causal interpretation could be viewed as a representation of working memory (for those familiar with psychological theories) or as a blackboard.

Figure 3 illustrates a causal interpretation. In the figure, an agent has a single goal (affiliation) that is threatened by the recent departure of a friend (the past "friend departs" action has one effect that deletes the "affiliation" state). This goal might be re-achieved if the agent joins a club. Appraisal assesses each case where an act facilitates or inhibits a fluent in the causal interpretation. In the figure, the interpretation encodes two "events," the threat to the currently satisfied goal of affiliation, and the potential re-establishment of affiliation in the future.

Each event is appraised along several appraisal variables by domain-independent functions that examine the syntactic structure of the causal interpretation:

- Perspective: from whose viewpoint is the event judged
- Desirability: what is the utility of the event if it comes to pass, from the perspective taken (e.g., does it causally advance or inhibit a state of some utility)
- Likelihood: how probable is the outcome of the event
- Causal attribution: who deserves credit or blame
- Temporal status: is this past, present, or future
- Controllability: can the outcome be altered by actions under control of the agent whose perspective is taken
- Changeability: can the outcome be altered by some other causal agent

Each appraised event is mapped into an emotion instance of some type and intensity, following the scheme proposed by Ortony et al [18]. A simple activation-based focus of attention model computes a current emotional state based on most-recently accessed emotion instances.

Coping determines how one responds to the appraised significance of events. Coping strategies are proposed maintain desirable or overturn undesirable in-focus emotion instances. Coping strategies essentially work in the reverse direction of appraisal, identifying the precursors of emotion in the causal interpretation that should be maintained or altered (e.g., beliefs, desires, intentions, expectations). Strategies include:

- Action: select an action for execution
- Planning: form intention to perform an act (the planner uses such intentions to drive its plan generation)
- Seek instrumental support: ask someone that is in control of an outcome for help
- Procrastination: wait for an external event to change the current circumstances

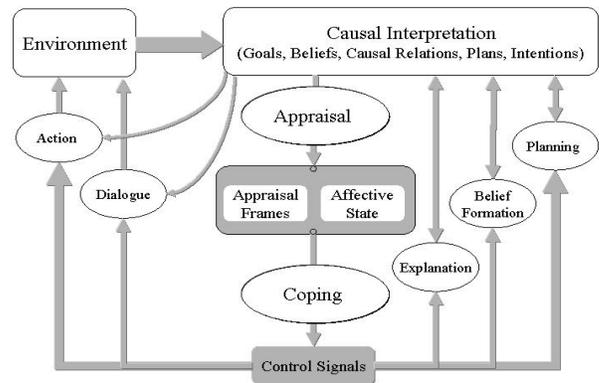


Figure 2: A reinterpretation of Smith and Lazarus

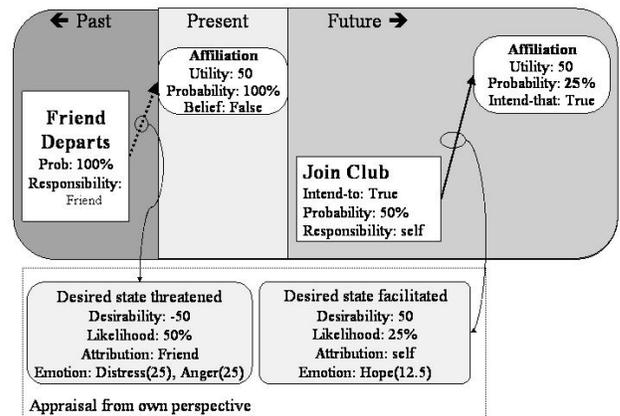


Figure 3: An example causal interpretation

- Positive reinterpretation: increase utility of positive side-effect of an act with a negative outcome
- Acceptance: drop a threatened intention
- Denial: lower the probability of a pending undesirable outcome
- Mental disengagement: lower utility of desired state
- Shift blame: shift responsibility for an action toward some other agent
- Seek/suppress information: form a positive or negative intention to monitor some pending or unknown state

Strategies give input to the cognitive processes that actually execute these directives. For example, planful coping will generate in intention to perform the join "join club" action, which in turn leads to the planning system to generate and execute a valid plan to accomplish this act. Alternatively, coping strategies might abandon the goal, lower the goal's importance, or re-assess who is to blame.

Not every strategy applies to a given stressor (e.g., an agent cannot engage in problem directed coping if it is unaware of an action that impacts the situation), however multiple strategies can apply. EMA proposes these in parallel but adopts strategies sequentially. EMA adopts a

small set of search control rules to resolve ties. In particular, EMA prefers problem-directed strategies if control is appraised as high (take action, plan, seek information), procrastination if changeability is high, and emotion-focus strategies if control and changeability is low.

In developing EMA's model of coping, we have moved away from the broad distinctions of problem-focused and emotion-focused strategies. Formally representing coping requires a certain crispness that is otherwise lacking in the problem-focused/emotion-focused distinction. In particular, much of what counts as problem-focused coping in the clinical literature is really inner-directed in an emotion-focused sense. For example, one might form an intention to achieve a desired state – and feel better as a consequence – without ever acting on the intention. Thus, by performing cognitive acts like planning, one can improve one's interpretation of circumstances without actually changing the physical environment.

### Related Work

EMA relates to a number of past appraisal models of emotion. Although we are perhaps the first to provide an integrated account of coping, EMA also contributes to the evolution towards domain-independent appraisal models. Early appraisal models focused on the mapping between appraisal variables and behavior and largely ignored how these variables might be derived. For example, Elliott's [19] Affective Reasoner, based on the OCC model [18], required a number of domain specific rules to appraise events (e.g., a goal at a football match is desirable if the agent favors the team that scored). More recent approaches have moved toward more abstract reasoning frameworks, largely building on traditional artificial intelligence techniques. For example, El Nasr and colleagues [20] use markov-decision processes (MDP) to provide a very general framework for characterizing the desirability of actions and events, including indirect consequences, but it can only represent a relatively small number of state transitions and assumes fixed goals. The closest approach to what we propose here is WILL [21] that ties appraisal variables to an explicit model of plans (which capture the causal relationships between actions and effects), although WILL, also, does not address the issue of blame/credit attributions, or how coping might alter this interpretation. We build on these prior models, extending them to provide better characterizations of causality and the subjective nature of appraisal that facilitates coping.

Prior computational work that has modeled the motivational function of emotions has largely focused on using emotion or appraisal to guide action selection. EMA appears to be the first to model the wider range of human coping strategies such as positive reinterpretation, denial, acceptance, shift blame, etc that alter beliefs, goals, etc.

Few computational models of emotion have been formally evaluated and most evaluations have focused on

external behaviors driven by the model rather than directly assessing aspects the emotion process. For example, most evaluations consider the interpretation of external behavior (e.g., are the behaviors believable?). More sophisticated work in this vein has tested more specific effects. For example, Prendenger [22] considered the impact of emotional displays on user stress and confidence and Lester evaluated the impact of emotional feedback on student learning. Additionally, there is now a sizable body of work on the impact of virtual human non-verbal behavior *in general* on human observers (e.g., [23]).

A small number of studies have tried to evaluate internal characteristics of an emotion process model. For example, Scheutz illustrated that the inclusion of an emotion process led artificial agents to make more adaptive decisions in a biologically inspired foraging task. We are unaware of any work, other than the work presented here, that has directly compared the dynamic processes of an emotion model against human data.

### Cognitive Dynamics

A key question for our model concerns its “process validity”: does EMA capture the unfolding dynamics of appraisal and coping. Rather than using an abstract overall assessment, such as observer self-reports of believability, we attempt to validate the model at a finer level. Ideally, we would like to show that the model faithfully captures how an arbitrary individual appraises a situation, how they cope, and how these appraisals and coping strategies evolve in response to changes in the situation. As a start, we address the simpler problem of how a “typical” person would respond emotionally to an evolving situation.

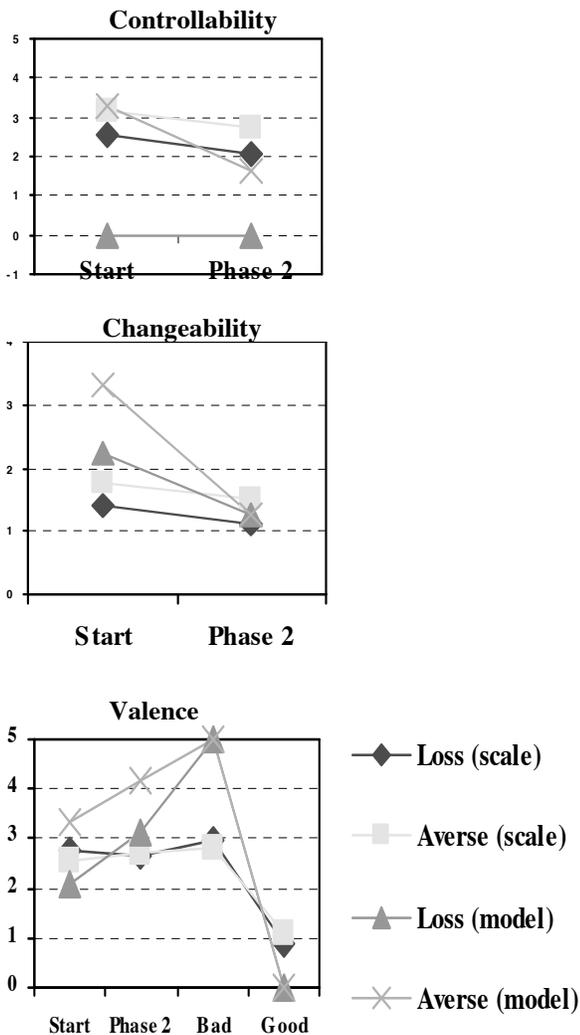
The Stress and Coping Process Questionnaire (SCPQ) [24] is used to assess a person's coping responses against a model of healthy adult behavior. A subject is presented a stereotypical situation, such as an argument with their boss. They are told to imagine themselves in that situation and queried on how they would feel, how they assess certain appraisal variables and what strategies they would use to cope. They are then given subsequent updates to the situation (e.g., some time has passed and the situation has not improved) and asked how their emotions/coping would dynamically unfold in light of systematic variations in both expectations and perceived sense of control. Based on their evolving pattern of responses, subjects are scored as to how closely their reactions correspond to a validated profile on how normal healthy adults respond.

In particular, the questionnaire describes two abstract situation conditions, each evolving over three discrete phases: an initial state, a state where some time passes without change, and ending with either a good or bad outcome. The “loss” condition presents a situation where a loss is looming in the future, the loss continues to loom for some time, and then the loss either occurs or is averted. In the “aversive” condition, some loss has oc-

curred but there is some potential to reverse it. After some time the loss is either reversed or the attempt to reverse it fails. The vocabulary used to describe these conditions is adjusted to produce a greater sense of control/changeability in the aversive condition.

From the SCPQ, normal subjects should illustrate the following trends:

- 1.1 Aversive condition should yield appraisals of higher controllability and changeability than the loss condition. (This effect was designed into the stimuli and its inclusion here is only validate it was achieved.)
- 1.2 Appraisal of controllability and changeability decrease over phases.
- 1.3 Negative valence should increase over phases and there should be a strong difference in valence on negative vs. positive outcomes.
- 1.4 Aversive condition should lead to more anger and



less sadness (the developers of the scale claim that this follows from the lack of appraised control in the loss condition).

- 2.1 Less appraised control should lead to less problem-directed coping
- 2.2 Less appraised control may produce more passivity
- 3.1 Lower ambiguity should produce a more limited the search for information
- 3.2 Lower ambiguity should yield more suppression of information about stressor
- 4 Less appraised control should produce more emotion-focused coping

### Methodology

We encode the situations as causal theories, evolve the situations in according to the SCPQ, and compare the model's appraisals and coping strategies to the trends indicated by the scale. Different phases are representing by changing the perceived likelihood of future outcomes. The SCPQ specifies the causal structure of the scenarios but we must set two parameters to complete each model, specifically the subjective probability of future actions in each phase and the utility of action outcomes.

Figure 3 illustrates the initial phase of the domain used for the aversive condition: some other agent performed an act that violated a desired state. Specifically in Figure 3, a friend is leaving which impacts the person's goal of affiliation (friendship), but a potential action under the control of the agent could lead to the desired outcome. Here that action is "join a club." In subsequent phases, we control the subjective probability that the future action will succeed/fail to be consistent with trend 1.1. In the aversive condition, the future action has 66% chance of succeeding, this drops to 33% in phase two, and in phase three is set to either zero or 100% percent, depending on if the bad or good outcome is modeled. The violated goal is given a high positive utility (100).

In the loss condition, the desired state is initially true and a future action under some other agent's control may defeat the goal. Again, probability across phases is adjusted. The chance of the loss succeeding is initially 50%, raises to 75% in phase two, and then is set to either 100% or 0%, depending on if the bad or good outcome is modeled. The threatened goal is given a high positive utility (100).

Some features of EMA do not map directly to terms in the SCPQ and were reinterpreted. We do not currently model ambiguity as an explicit appraisal variable. Since the only ambiguity in the SCPQ scenarios relates to the success of pending outcomes, we equate ambiguity with changeability for the purposes of this evaluation. Following our use of the OCC mapping of appraisal variables to emotion types, EMA also does not directly appraise "sadness" but

rather derives “distress” (an undesired outcome has occurred). For this evaluation we equate “sadness” with “distress.” Finally, trend 1.3 depends on an overall measure of “valence” that EMA does not support. Given that we appraise individual events and an event may have good and bad aspects, for the purpose of this evaluation we derive an aggregate valence measure that sums the intensities of undesirable appraisals and subtracts from the intensities of positive appraisals.

### Results

Trends 1.1 and 1.3 are fully supported by the model. Trend 1.2 is partially supported. The appraisal of controllability and changeability decrease over phases in the aversive condition, but only the changeability decreases over the loss condition as the model determines there is no control over the looming loss. Trend 1.4 is partially supported. There is more anger in the aversive condition, however there is also more sadness, contrary to the prediction. Rather than having higher sadness, the loss condition yielded only fear until the bad outcome (where fear becomes sadness).

Trends 2.1 and 2.2 are supported. In the aversive condition, the model forms an intention to restore the loss only when its probability of success is high (phase 1). In the loss condition, no action in the causal interpretation can influence the pending loss so control is low and no problem-directed strategies are selected. When changeability is high (phase 1 of both conditions), the model suggests a wait-and-see strategy, which is rejected in later phases.

Trends 3.1 and 3.2 are fully supported. When changeability is high, the model proposes monitoring the truth-value of the state predicate that has high probability of changing. As changeability drops, the model proposes strategies that suppress the monitoring of these states.

Trend 4 is supported. As the control drops, proposed strategies tend towards emotion-focused (see Table 1).

<b>Table 1</b>	Aversive	Loss
Phase 1	Seek information Take action	Suppress information Procrastinate Seek inst. support
Phase 2	Distance Suppress info.	Distance Suppress information Resignation Wishful thinking
Good	Accept responsibility	
Bad	Distance Suppress info.	Distance Suppress information

### Discussion

The model supports most of the trends predicted by SCPQ. Two departures deserve further mention. The loss

condition should have produced more sadness than the aversive condition but the opposite occurred. This indicates that the OCC model’s definition of “distress” is inappropriate to model sadness. In fact, many models tie the attribution of sadness directly to the perceived sense of control over the situation (e.g., LAZARUS), and this can be straightforwardly added to EMA.

A second departure from the human data is that EMA attributes zero control to the agent in the loss condition. This is due to the fact that, in our encoding, the only action that could impact the goal is the “impending loss” task, which is under control of the other agent. This is clearly too strong and could be relaxed by adding some other task to the domain model under the agent’s control that could influence the likelihood of the loss.

There are pros and cons to our current methodology from the standpoint of evaluation. On the plus side, the situations in the instrument were constructed by someone outside our research group, and thus constitute a fairer test of the approach’s generality than what is often performed (though we are clearly subject to bias in our selection of a particular instrument). Further, by formalizing an evolving situation, this instrument directly assesses the question of emotional dynamics, rather than single situation-response pairs typically considered in evaluations. On the negative side, the scenarios were described abstractly and we had considerable freedom in how we encoded the situations into a causal mode.

A more general concern is the use of aggregate measures of human behavior. People show considerable individual difference in their appraisal and coping strategy. In this evaluation, however, we compare the model to aggregate trends that may not well-approximate any given individual. This concern is somewhat mitigated by the fact that the SCPQ scale is intended to characterize individuals in terms of the “normalcy” of their emotional behavior and has been validated for this use. However, a more rigorous test would be to fit to individual reports based on their perceived utility and expectations about certain outcomes.

The exercise of encoding situations into a domain theory acceptable by EMA clearly delineates its current limits. For example, the model does not appraise ambiguity.

### Behavioral Impact

The second study evaluates whether expressive behaviors influence human subjects in the same way they would be influenced by the emotional behavior of other humans. Social psychologists have identified a number of interpersonal functions of emotional behavior in human-to-human interactions. In moving beyond the concept of believability, we ask if these specific functions can play a similar role in agent-to-human interactions.

Here we focus on the phenomena of social referencing [25], in which a person faced with an ambiguous situation

is influenced by the emotional/evaluative reactions of others. The question here is if a virtual human can similarly influence a person's decision-making.

Subjects view one of two pre-recorded clips taken from our Mission Rehearsal Exercise training simulation. The clips are presented on a 8'x10' screen in a small theater setting. The sequence portrays a dialogue between a U.S. Army lieutenant (who's disembodied voice is heard from behind and his platoon sergeant (a subordinate) standing in the center of the screen. A medic under the lieutenant's command is kneeling beside an injured child, and several soldiers in the platoon stand near by. The onscreen characters are virtual humans (and readily identified as human facsimiles). Subjects are told to pay attention to the discussion between the lieutenant and the platoon sergeant. Specifically, they are told these two characters will be discussing options about how to balance two needs: On the one hand, they must remain on the scene to help the injured child and avoid splitting their forces (the SGT preferred condition). On the other hand, they must provide reinforcements to another platoon (designated Eagle-1-6) several miles down the road.

We hand-edited a recording of an actual simulation run two create two conditions that differ only in terms of the behavior of the bystander soldiers (the behaviors and voice of the sergeant and the voice of the lieutenant are constant across the clips). In the SGT-Congruent condition, the soldiers display head nods and facial expressions that express agreement with the sergeant, and disagreement with the lieutenant. In the SGT-Incongruent condition, the soldiers display head nods and facial expressions that express disagreement with the sergeant and agreement with the lieutenant.

#### Methodology

This pilot study used 19 volunteers (17 male and 2 female) who were rewarded with pizza for their participation. Subjects were divided into two groups and each group saw one clip, filled out a questionnaire, then saw the other clip and filled out the same questionnaire. The order of the presentation was reversed between groups. Subjects were asked what they thought was the best decision, how confident they were and why. Subjects were also asked what decision they believed the soldiers preferred, what emotions were expressed by the sergeant, and how natural the interaction seemed. Confidence and Naturalness were assessed on a seven point scale.

#### Results

Table 2 summarizes the key results.

<b>Table 2</b>	SGT Congruent	SGT Incongruent
Subject agrees with SGT	78%	33%
Confidence	4.2	4.2
Soldiers agree with SGT	100%	6%
Confidence	6.2	5.8
Naturalness	3.5	2.8

The main hypothesis was supported. Subjects' decisions were significantly influenced by the non-verbal behavior of the soldiers in the predicted direction ( $p=0.018$ ). In a third of those cases, subjects explicitly noted the agreement of the soldiers as their justification, however many subjects (28%) justified their decision in terms of elements of the task. For example, several subjects in the SGT-incongruent condition stated that the current location seemed safe enough to divide the forces. There was a nearly significant trend toward viewing the incongruent condition as less natural ( $p=0.19$ ). Many subjects reported that the behaviors in this condition seemed too synchronized, which may have contributed to this trend

Subjects were asked to assess the expressed emotions of the sergeant but also what they imagined his emotions to be. Subjects gave almost identical reports of observed emotion across condition (modest amount each of anger, distress and anxiety as the most expressed emotions). There was a non-significant trend towards observing more fear in the incongruent conditions (2.4 vs. 1.3) even though the behavior was the same.

#### Discussion

Subjects' decisions were strongly influenced by the non-verbal behavior of bystanders in the direction predicted by social referencing. This suggests that non-verbal behavior of virtual characters can generate similar interpersonal influence to what people have on each other.

Whether this influence occurred through the same mechanism as social referencing, however, remains an open question. For example, imagine that instead of using non-verbal behavior, we simply added the words "the lieutenant is right" at the top of the screen. This may have had the same effect but it is difficult to argue that it occurs through social referencing (unless the reference set is extended to include the scientists running the exercise). Indeed, if the goal is simply to influence the observer's decision, we are unclear how to distinguish between these two manipulations. If social referencing were at play, then presumably the effect would be modified by manipulations that increased or decreased the presumed expertise or bias of the bystanders.

Again, we see distinct advantages of this evaluation. By directly assessing the influence of emotional behaviors on human performance variables, we can be more concrete on the utility of emotion in agent-human interaction. For

example, social referencing is important for applications that either need to influence human choice, or wish to educate people about such influences.

## Summary

Spurred by a range of potential applications, there has been a growing body of research in computational models of human emotion. To advance the development of these models, it is critical that we begin to contrast them against the phenomena they purport to model.

We have presented two approaches to evaluate an emotion model. In our process comparison we compared internal attributions and process dynamics against human data, using a standard clinical instrument. Remarkably, the model did quite well. And, as expected, the comparison helped identify where the model needed further development. In our "impact" evaluation, we compared the impact of non-verbal behaviors against standard psychological phenomena, again with positive results.

As with any new discipline, evaluation of affective systems has lagged far behind advances in computation models. This situation is slowly changing as a number of groups move beyond simple metrics and move toward more differentiated notions of the form and function of expressed behavior (e.g. [22, 26]). This paper contributes to this evolution.

## Acknowledgements

This work was funded by the Department of the Army under contract DAAD 19-99-D-0046. Lilia Moshkina produced the stimuli used the social referencing experiment. Any opinions, findings, and conclusions expressed in this article are those of the authors and do not necessarily reflect the views of the Department of the Army.

- [1] G. L. Clore and K. Gasper, "Feeling is believing: Some affective influences on belief," in *Emotions and Beliefs: How Feelings Influence Thoughts, Studies in Emotion and Social Interaction: Second Series*, N. Frijda, A. S. R. Manstead, and S. Bem, Eds. Paris: Cambridge University Press, 2000, pp. 10-44.
- [2] A. R. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon Books, 1994.
- [3] R. J. Davidson, K. Scherer, and H. H. Goldsmith, "Handbook of Affective Sciences," in *Series in Affective Science*, R. J. Davidson, P. Ekman, and K. Scherer, Eds. New York: Oxford University Press, 2003.
- [4] L. Cosmides and J. Tooby, "Evolutionary Psychology and the Emotions," in *Handbook of Emotion*, M. Lewis and J. Haviland-Jones, Eds., Second Edition ed. NY, NY: Guilford Press, 2000, pp. 91-115.
- [5] J. Gratch, "Émile: marshalling passions in training and education," presented at Fourth International Conference on Intelligent Agents, Barcelona, Spain, 2000.
- [6] J. Gratch and S. Marsella, "Tears and Fears: Modeling Emotions and Emotional Behaviors in Synthetic Agents," presented at Fifth International Conference on Autonomous Agents, Montreal, Canada, 2001.

- [7] S. Marsella and J. Gratch, "Modeling coping behaviors in virtual humans: Don't worry, be happy," presented at Second International Joint Conference on Autonomous Agents and Multi-agent Systems, Melbourne, Australia, 2003.
- [8] J. Rickel, S. Marsella, J. Gratch, R. Hill, D. Traum, and W. Swartout, "Toward a New Generation of Virtual Humans for Interactive Experiences," in *IEEE Intelligent Systems*, vol. July/August, 2002, pp. 32-38.
- [9] S. Marsella, W. L. Johnson, and C. LaBore, "Interactive Pedagogical Drama," presented at 4th International Conference on Autonomous Agents, Montreal, Canada, 2000.
- [10] K. R. Scherer, A. Schorr, and T. Johnstone, "Appraisal Processes in Emotion," in *Affective Science*, R. J. Davidson, P. Ekman, and K. R. Scherer, Eds.: Oxford University Press, 2001.
- [11] J. Velásquez, "When robots weep: emotional memories and decision-making.," presented at Fifteenth National Conference on Artificial Intelligence, Madison, WI, 1998.
- [12] E. Peacock and P. Wong, "The stress appraisal measure (SAM): A multidimensional approach to cognitive appraisal," *Stress Medicine*, vol. 6, pp. 227-236, 1990.
- [13] R. Lazarus, *Emotion and Adaptation*. NY: Oxford University Press, 1991.
- [14] K. Scherer, "On the nature and function of emotion: A component process approach," in *Approaches to emotion*, K. R. Scherer and P. Ekman, Eds., 1984, pp. 293-317.
- [15] J. Blythe, "Decision Theoretic Planning," in *AI Magazine*, vol. 20(2), 1999, pp. 37-54.
- [16] M. Bratman, "What is intention?," in *Intentions in Communication*, P. Cohen, J. Morgan, and M. Pollack, Eds. Cambridge, MA: MIT Press, 1990.
- [17] B. Grosz and S. Kraus, "Collaborative Plans for Complex Group Action," *Artificial Intelligence*, vol. 86, 1996.
- [18] A. Ortony, G. Clore, and A. Collins, *The Cognitive Structure of Emotions*: Cambridge University Press., 1988.
- [19] C. Elliott, "The affective reasoner: A process model of emotions in a multi-agent system," Northwestern University Institute for the Learning Sciences, Northwestern, IL, Ph.D Dissertation 32, 1992.
- [20] M. S. El Nasr, J. Yen, and T. Ioerger, "FLAME: Fuzzy Logic Adaptive Model of Emotions," *Autonomous Agents and Multi-Agent Systems*, vol. 3, pp. 219-257, 2000.
- [21] D. Moffat and N. Frijda, "Where there's a Will there's an agent," presented at Workshop on Agent Theories, Architectures and Languages, 1995.
- [22] H. Prendinger, S. Mayer, J. Mori, and M. Ishizuka, "Persona Effect Revisited," presented at Intelligent Virtual Agents, Kloster Irsee, Germany, 2003.
- [23] N. C. Kramer, B. Tietz, and G. Bente, "Effects of embodied interface agents and their gestural activity," presented at Intelligent Virtual Agents, Kloster Irsee, Germany, 2003.
- [24] M. Perrez and M. Reicherts, *Stress, Coping, and Health*. Seattle, WA: Hogrefe and Huber Publishers, 1992.
- [25] J. J. Campos, "The importance of affective communication in social referencing: a commentary on feinman," *Merrill-Palmer Quarterly*, vol. 29, pp. 83-87, 1983.
- [26] A. Cowell and K. M. Stanney, "Embodiment and Interaction Guidelines for Designing Credible, Trustworthy Embodied Conversational Agents," presented at Intelligent Virtual Agents, Kloster Irsee, Germany, 2003.