TOWARDS A VALIDATED MODEL OF THE INFLUENCE OF EMOTION ON HUMAN PERFORMANCE

Jonathan Gratch*
University of Southern California Institute for Creative Technologies
13274 Fiji Way, Marina del Rey, CA 90292

Stacy Marsella
University of Southern California Information Sciences Institute
4676 Admiralty Way, Marina del Rey, CA 90292

1. INTRODUCTION

Emotions play a powerful, central role in our lives and not surprisingly, they play an equally central role in military planning and training. Emotions shape how we perceive the world, bias our beliefs, influence our decisions and in large measure guide how we adapt our behavior to the physical and social environment. Though advances in psychology and neurophysiology have highlighted the rational and adaptive nature of some emotional responses (Damasio, 1994; Lazarus, 1991), clearly emotions can be influenced and exploited as a social tool and this is the essence of their value to military operations. The ancient Greeks wrote about the rhetorical power of pathos, an appeal to emotion, and military planners throughout history have incorporated an emotional element into their military doctrine. Machiavelli wrote that to motivate citizens to withstand a long siege one should encourage “fear of the cruelty of the enemy” (Machiavelli, 1998). The following quote from a U.S. Army leadership manual illustrates the role of emotions in more operational terms (Leaders’ manual for combat stress control, 1994):

Commanders, while shielding their own troops from stress, should attempt to promote terror and disintegration in the opposing force. … Some examples of stress-creating actions are attacks on his command structure; the use of artillery, air delivered weapons, smoke; deception; psychological warfare; and the use of special operations forces. Such stress-creating actions can hasten the destruction of the enemy’s capability for combat.

With such explicit proscriptions, it is potentially troubling that emotions and other psychological factors are so poorly modeled by the computer simulations that increasingly inform and shape military operations. The acknowledged weakness of simulation technology is its failure to capture the essence of human behavior (Pew & Mavor, 1998). The field of artificial intelligence has made great strides in producing algorithms that plan, act and learn from experience; however these techniques have grown out of a narrowly rational conception of intelligent behavior. Contemporary artificial intelligence approaches arose from normative perspectives on intelligence such as decision theory, logical deduction, and game theory. Though rationality seems a reasonable goal for the engineering applications that have motivated artificial intelligence, these models have significant shortcomings when it comes to modeling human behavior. In contrast, cognitive modeling approaches that explicitly capture human capabilities and limitations have tended to focus on narrow scientific phenomena – explaining reaction time data or the impact of priming on recall tasks – and are less appropriate for modeling the broad reasoning capabilities demanded by modeling and simulation applications. The consequence is that modeling and simulation systems are particularly ill-suited for capturing the influence that factors such as stress and emotion can have on military outcomes. The risk is that such systems implicitly institutionalize a misleading view of human behavior; a view that increasingly shapes military training, planning, and acquisition decisions.
In this article, we provide an overview of our recent progress in developing and validating human behavior models that incorporate the influence of emotion on cognition. This work has been developed over the last four years within the context of Mission Rehearsal Exercise (MRE) project (see Figure 1), a research prototype designed to support immersive leadership training. MRE puts trainees into unscripted human-oriented simulations, where they can improvise solutions with virtual humans (Gratch et al., 2002). These software entities look and act like people and can engage in conversation and collaborative tasks, but unlike robots, they exist in simulated environments. The technology underlying virtual humans is a natural, albeit more ambitious extension of the approaches used to model human decision-makers in military simulations (Hill, Chen, Gratch, Rosenbloom, & Tambe, 1997). Virtual humans must act and react in their simulated environment, drawing on the disciplines of automated reasoning and planning. To hold a conversation, they must exploit the full gamut of natural language research, from speech recognition and natural language understanding to natural language generation and speech synthesis. Providing human bodies that can be controlled in real time delves into computer graphics and animation. And because a virtual human looks like a human, people expect it to behave like one as well and will be disturbed by, or misinterpret, discrepancies from human norms. Thus, virtual human research must draw heavily on psychology and communication theory to appropriately convey nonverbal behavior, emotion, and personality.

2. MODELING EMOTION

To model emotion in simulation, we need to consider what emotion is, how it impacts human behavior and how it can be computationally modeled within a simulation environment. In the simplest view, emotions can be viewed and studied as just simple patterned behavioral and physiological responses to specific stimuli. However, increasingly, the neural and psychological research has argued that emotion is more than simple patterned response and in fact that there is a tight integration of emotion and cognitive processes. In our work, we draw heavily on appraisal theories of emotion that claim emotions can arise from a cognitive assessment of the environment and that this assessment in turn influences behavior (Arnold, 1960; Lazarus, 1991; K. R. Scherer, Schorr, & Johnstone, 2001). In appraisal theories (see Figure 2) emotions are part of an adaptive, flexible responses to the environment. This flexible response is realized by two basic processes: appraisal and coping (Lazarus, 1991). Appraisal generates emotion by a cognitive assessment of the person-environment relationship along several key dimensions, including: whether an event facilitated or inhibited the person’s goals; how critical was the impact of this event; who deserves blame or credit. Coping is the process by which the person deals with emotion. Two classes of emotion have been identified, problem-focused coping and emotion-focused coping. Individuals that adopt problem-focused strategies act externally on the world to address the factors leading to emotion. For example, if some event threatens a person’s goals, leading to anger, the person may take action to counter that threat. Individuals that adopt emotion-focused coping act internally to change beliefs or attention. For example, a person may deny the threat is real, be resigned to the fact that the threatened goal cannot be achieved or in some way try to avoid thinking about the threat. Within these broad classes of coping, people cope with emotions in myriad ways and psychologists have documented a rich set of coping strategies. Different individuals tend to adopt stable and characteristic “coping styles” that are correlated with personality. Furthermore, coping and appraisal interact and unfold over time, leading to dynamic and characteristic changes in emotional state that has been noted by several emotion researchers (Lazarus, 1991; K. Scherer, 1984): a person may “feel” distress for an event (appraisal), which motivates the shifting of blame (coping), which leads to anger (re-appraisal).

EMA is a computational model based on appraisal theory and described in detail elsewhere (Gratch & Marsella, 2001, 2004; Marsella & Gratch, 2003). 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
In terms, we have found it most natural to build on the causal representations developed for decision-theoretic planning (e.g., (Blythe, 1999)) and augment them with methods that explicitly model commitments to beliefs and intentions (Grosz & Kraus, 1996). 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). The decision-theoretic concepts of utility and probability are key for modeling appraisal variables of desirability and likelihood. Explicit representations of intentions and beliefs are critical for properly reasoning about causal attributions, as these involve reasoning if the causal agent intended or foresaw the consequences of their actions (Shaver, 1985). As we will see, commitments to beliefs and intentions also play a role in modeling coping strategies.

In EMA, the agent’s interpretation of its “agent-environment relationship” is reified in an explicit representation of beliefs, desires, intentions, plans and probabilities. Following a blackboard-style model, this representation (corresponding to the agent’s working memory) encodes the input, intermediate results and output of reasoning process that mediate between the agent’s goals and its physical and social environment (e.g., perception, planning, explanation, and natural language processing). 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, the causal interpretation represents the agent’s current view of the agent-environment relationship, which changes with further observation or inference. We treat appraisal as a set of feature detectors that map features of this representation into appraisal variables. For example, an effect that threatens a desired goal is assessed as a potential undesirable event. Coping sends control signals to auxiliary reasoning modules (i.e., planning, belief updates, etc.) to overturn or maintain those features that yielded the appraisals. For example, coping may resign the agent to the threat by abandoning the desired goal.

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 goal 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 (Ortony, Clore, & Collins, 1988). 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 an intention to perform some act (the planner uses 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
- 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, going back to Figure 3, planful coping will generate in
intention to perform the "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 a computational 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 lacking from 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 a 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 ones interpretation of circumstances without actually changing the physical environment.

### 3. EVALUATION

A key question for our model concerns its “process validity”: does the model capture the unfolding dynamics of appraisal and coping. Rather than using an abstract overall assessment, such as observer self-reports of “believability,” we would like to directly compare the internal variables of the model to human data, assessing emotional responses, but also the value of appraisal variables, coping tendencies, and in particular, how these assessments change in response to an evolving situation.

Although human mental processes cannot be observed directly, several clinical instruments have been developed to assess this information indirectly through interactive questionnaires. For example, the Stress and Coping Process Questionnaire (SCPQ) (Perrez & Reicherts, 1992) is a clinical instrument used to assess a human subject’s coping process against an empirical model of normal, healthy adult behavior. A subject is presented a stereotypical episode and their responses are measured several times as the episode evolves. For example, they are told to imagine themselves in an argument with their boss and are queried on how they would feel (emotional response), how they appraise certain aspects of the current situation (appraisal variables) and what strategies they would use to confront the situation (coping strategies). They are then presented updates to the situation (e.g., they are told some time has passed and the situation has not improved) and asked how their emotions/coping would dynamically unfold in light of these manipulations. The episodes are evolved systematically to alter 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.

Using such a scale has the advantage that it provides an independently derived corpus of evolving situations and a ready source of human data, though it does not provide data on individual differences. Ideally, we would like to show that EMA captures how an arbitrary individual appraises a situation given knowledge of their initial beliefs and preferences, or at least models the most common response. As a start however, and given the practical difficulties in obtaining individual information, we compare EMA against aggregate data from the SCPQ. This instrument averages observations across multiple subjects and attempts to characterize “typical” human responses. Given the variability of human emotional behavior, we believe it is important to start by comparing against such normalized responses.

Figure 4 illustrates one of the episodes from the SCPQ.
all are generated from a grammar that encodes two prototypical stressful episodes. Episodes evolve over three discrete phases: an initial state, a state where some time passes without change, and an ending phase which can either result in a good or bad conclusion. The loss condition prototype presents an episode where some 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 prototype, some bad outcome has occurred but there is some potential to reverse it. After some time the undesirable outcome is either reversed or the attempt to reverse it fails. In all, there are four canonical situations (loss-good, loss-bad, aversive-good and aversive-bad) each of which are represented by multiple variants in the scale. The aversive condition is designed to convey a greater sense of control/changeability, and the vocabulary is selected and empirically validated to produce this effect. Figure 4 illustrates a loss-bad episode.

When used as a diagnostic tool, a patient would fill out their interpretation of the set of evolving situations. These are scored with respect to how closely they follow the trends exhibited by healthy adults. These trends include:

1.1 Aversive condition should yield appraisals of higher controllability and changeability than the loss condition (this follows from the design of the stimuli)
1.2 Appraisal of controllability and changeability decrease over phases (as likelihood of change drops)
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 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

Our intention is to use the scale as a diagnostic instrument to ascertain if the judgments made by our model fall with the expected range of responses of normal healthy adults. Rather than attempting to parse English and use the scale directly, we take advantage of the fact that all of the episodes in the scale correspond to one of four canonical scenarios. Thus, we encode the causal structure of these four episodes into EMA.

Methodology

We encode the four canonical episodes in the SCPQ as evolving causal theories and compare the model’s appraisals and coping strategies to the trends indicated by the scale. Consistent with how the SCPQ is used, we allow the model to propose coping strategies, but these proposals do not influence subsequent phases (the model proposes strategies but their effects are preempted). The evolution of each episode is encoded by changing the perceived likelihood of future outcomes at each phase in the episode. The SCPQ provides the basic causal structure of episodes 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: an action executed by some other agent in the past (friend leaving) makes false some desired state (friendship), but there is some potential action under the control of the agent with no preconditions and one effect that could lead to the desired outcome (join a club). (Labels on states and actions do not impact the model.) In subsequent phases, we alter the subjective probability that the future action will succeed/fail. In the aversive condition, the future action has 66% chance of succeeding, this drops to 33% in phase two, and then is set to either zero or 100% percent, depending on if the bad or good outcome is modeled. The violated goal has high positive utility (100).

Figure 5 illustrates the initial phase of the domain for the loss condition: a desired state is initially true and a future action potentially executed by another agent may make this state false. In subsequent phases, we alter the subjective probability that the future action will succeed/fail. In the aversive condition, the future action has 66% chance of succeeding, this drops to 33% in phase two, and then is set to either zero or 100% percent, depending on if the bad or good outcome is modeled. The violated goal has high positive utility (100).

Figure 5 illustrates the initial phase of the domain for the loss condition: a desired state is initially true and a future action potentially executed by another agent may make this state false. 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

\footnote{SCPQ treats this as two distinct sub-trends, distinguishing between two types of emotion-directed strategies. As Smith and Lazarus do not make this distinction, we collapse them.}
modeled. The desired state has high positive utility (100).

Some terms used in the SCPQ do not map directly to representational primitives in EMA and had to be reinterpreted. EMA does 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. As EMA incorporates the OCC mapping of appraisal variables to emotion types (Ortony et al., 1988), our model 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 our model 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. We revisit some of these decisions in the discussion.

Results

Trends 1.1 and 1.3 are supported by the model: the aversive condition is appraised as more controllable and changeable and negative valence increases across phases in both conditions. Trend 1.2 is fully supported for the aversive condition but only partially supported in the loss condition: EMA correctly deduces that the situation is less likely to change across phases, but it decides that the agent has no control over the loss, even in phase 1. Trend 1.4 is also partially supported: there is more anger in the aversive condition, however these is also more sadness, contrary to the prediction. Rather than yielding higher sadness, EMA appraised only fear in the initial phases of the loss condition. Sadness arises only in the bad outcome, when the looming loss becomes certain.

Trends 2.1 and 2.2 are both 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 known action 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 the model finds the situation likely to improve on its own (high changeability), it 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). In the aversive condition, for example, EMA initially forms an intention to execute the “join a club” action (take action) and forms an intention to monitor the truth value of the desired state (seek information). As the likelihood that the action will succeed diminishes, the agent forms an intention to avoid monitoring the status of the desired state (suppress information) and begins to lower its attachment to the goal by lowering its utility (mental disengagement). This trend is reinforced in the bad outcome, but is reversed if the action succeeds (good outcome).

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 may indicates that the OCC model’s definition of “distress” is inappropriate for modeling sadness. OCC appraises distress whenever an undesirable event has occurred, however, many theories argue that the attribution of sadness is also related to the perceived sense of control over the situation (e.g., (Lazarus, 1991)). This alternative definition could be straightforwardly added to our model.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Aversive</th>
<th>Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1</td>
<td>Seek information</td>
<td>Suppress information</td>
</tr>
<tr>
<td></td>
<td>Take action</td>
<td>Procrastinate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Seek instr. support</td>
</tr>
<tr>
<td>Phase 2</td>
<td>Mental disengagement</td>
<td>Mental disengagement</td>
</tr>
<tr>
<td></td>
<td>Suppress information</td>
<td>Suppress information</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resignation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wishful thinking</td>
</tr>
<tr>
<td>Good</td>
<td>Accept responsibility</td>
<td></td>
</tr>
<tr>
<td>Bad</td>
<td>Mental disengagement</td>
<td>Mental disengagement</td>
</tr>
<tr>
<td></td>
<td>Suppress information</td>
<td>Suppress information</td>
</tr>
</tbody>
</table>
A second departure from the human data is that the model appraises zero control in the loss condition across all phases. This is due to the fact that, in our encoding, another agent is represented as the actor for the “looming loss” action, meaning the agent has no direct control and, as this action has no preconditions that could be confronted, there is no indirect control as well. This is clearly too strong and probably does not reflect the causal structure that people recover when they read the SCPQ episodes. This assumption could be relaxed by adding some other action to the domain model executable by the agent 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 some freedom in how we encoded the situations into a causal mode, potentially biasing our results.

A more general concern is the use of aggregate measures of human emotional 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 “normality” 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.

ACKNOWLEDGMENTS

This work was funded by the Department of the Army under contract DAAD 19-99-D-0046. 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.

REFERENCES

proceedings of the Fourteenth National Conference on Artificial Intelligence and the Ninth Conference on Innovative Applications of Artificial Intelligence (AAAI/IAAI97), Providence, RI.


