Two Sides of Appraisal: Implementing Appraisal and Its Consequences within a Cognitive Architecture

Eva Hudlicka

Psychometrix Associates, Inc.
1805 Azalea Drive, Blacksburg, VA , USA
evahud@earthlink.net

Abstract

Appraisal processes provide an affective assessment of an agent's current situation, in light of its needs and goals. This paper describes a computational model of the appraisal process, implemented within the broader context of a cognitive agent architecture. A particular focus here is on modeling the interacting influences of states and traits on perception and cognition, including their effects on the appraisal process itself. These effects are modeled by manipulating a series of architecture parameters, such as the speed and processing capacity of the individual modules. The paper presents results of an evaluation experiment modeling the behavior of three types of agents: 'normal', 'anxious', and 'aggressive'. The appraisal model generated different affective appraisals of the same set of external circumstances for the different agent types, resulting in distinct emotions, and eventually leading to observable differences in behavior. The paper concludes with a brief discussion of some of the issues encountered during the appraisal model development.

Introduction and Objectives

Appraisal is a core component of cognitive-affective processing. Its objective is to provide an affective assessment of a current situation, internal or external, in light of the organism's needs and goals. The processes comprising appraisal are subject to a variety of influences, including individual history, personality, and current affective state.

The past decade has witnessed a rapid growth in appraisal research, including progress towards resolving the primacy of cognition vs. affect debate (e.g., Scherer, 2003), collection of empirical data (Roseman, 2001), refinement of theories to be increasingly process-oriented (and thus lend themselves to implementation) (e.g., Smith & Kirby, 2001), and, to some extent, the construction of computational models (see reviews in Wehrle & Scherer, 2001; Hudlicka & Fellous, 1996)). Computational models represent an important tool in appraisal research, because the construction of models of hypothesized mechanisms requires a degree of operationalization which often reveals gaps in knowledge or theoretical inconsistencies. Computational modeling thus provides opportunities for validation, and for the generation of alternative hypotheses explaining specific data or phenomena.

The objective of this on-going effort is to develop a computational model of the appraisal process, implemented within the broader context of a cognitive agent architecture. The architecture controls the behavior of a simulated agent, within a simulated environment, by processing incoming stimuli and generating responses. Implementing the appraisal process within this context not only provides a dynamic set of stimuli, to which the agent must respond, but also the internal mental constructs necessary for the appraisal process - beliefs, goals, expectations. These constructs are dynamically generated in response to changing external and internal conditions, and enable the implementation of a simulated in vivo appraisal functioning. A particular focus here is on modeling the effects of states and traits on perception and cognition, including their effects on the appraisal process itself. Below we discuss the model structure, preliminary results, issues identified during the implementation, and future work.

Agent Architecture, Trait / State Modeling, and Affect Appraiser

We implemented a staged, multi-level appraisal model, within the context of a cognitive agent architecture (MAMID - Hudlicka, 2002). We first describe this architecture, to provide the necessary context for a discussion of the trait / state modeling methodology, and the appraisal model, which follow.

Cognitive Architecture

The cognitive architecture implements a sequential see-think-do processing sequence (figure 2-1), consisting of the following modules: *sensory pre-processing*, translating incoming data into task-relevant cues; *attention*, filtering incoming cues and selecting a subset for processing; *situation assessment*, integrating individual cues into an overall situation assessment; *expectation generation*, projecting current situation onto possible future states; *affect appraiser,* deriving the affective state (both valence and four of the basic emotions) from a variety of external and internal elicitors, both static and dynamic; *goal selection*, selecting critical goals for achievement; and *action selection,* selecting the best actions for goal achievement. These *modules* map the incoming stimuli (cues) onto the outgoing behavior (actions), via a series of intermediate internal representational structures (situations, expectations, and goals), collectively termed *mental constructs*. This mapping is enabled by long-term memories (LTM) associated with each module, represented in terms of belief nets or rules. *Mental constructs* are characterized by their attributes (e.g., familiarity, novelty, salience, threat level, valence, etc.), which influence their processing; that is, their rank and the consequent likelihood of being processed within a given architecture cycle; (e.g., cue will be attended, situation derived, goal or action selected). (Note that the availability of the mental constructs from previous frames of the execution cycle allows for dynamic feedback among constructs, and thus departs from a strictly sequential processing sequence.)

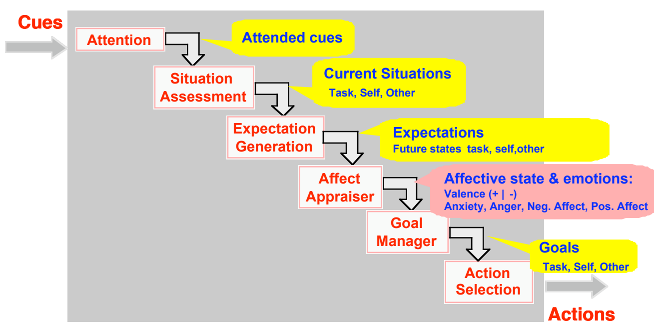

Figure 2-1: MAMID Cognitive Architecture: Modules & Mental Constructs

Modeling State and Trait Effects

The effects of states and traits on attention, perception, and decision-making have been extensively documented (e.g., Williams et al., 1997; Isen, 1993; Forgas, 2003). States and traits influence the speed, capacity and accuracy of attention and working memory; long-term memory structure, encoding, and recall; and interpretive, goal-management, and decision-making processes (figure 2-2). Some of the more robust findings regarding state effects include fear and anxiety-induced reduction in attention and working memory capacities, and attentional and interpretive threat-bias. Trait effects can be both structurally-oriented (e.g., higher proportion of self- and threat-oriented memory schemas associated with low emotional stability (Matthews & Deary, 1998)), and functionally-oriented (e.g., traits influence the dynamic characteristics of states such as onset triggers, ramp up and decay rates, and maximum intensities) (Eid, 2001)).

We used a previously described methodology for modeling state and trait effects within a cognitive architecture (Hudlicka, 2002; 1998), which consists of mapping particular state / trait profiles onto specific architecture parameter values (figure 2-3). These parameters then control processing within individual architecture modules.

Functions implementing these mappings were constructed on the basis of the available empirical data, as outlined above. For example, reduced attentional and working memory (WM) capacity, associated with anxiety and fear, are modeled by dynamically reducing the attentional and WM capacity of the architecture modules, which then reduces the number of constructs processed (fewer cues attended, situations derived, expectations generated, etc.). Attentional threat bias is modeled by higher ranking of threatening cues, thus increasing their likelihood of being attended, and by higher ranking of threatening situations and expectations, thus increasing the chances of a threatening situation / expectation being derived. Trait-linked structural differences in LTM are supported by allowing the flexible selection of alternative LTM clusters, reflecting distinct personality traits. Traits also influence the dynamic characteristics of the emotional responses (ramp up, decay, and maximum intensities).

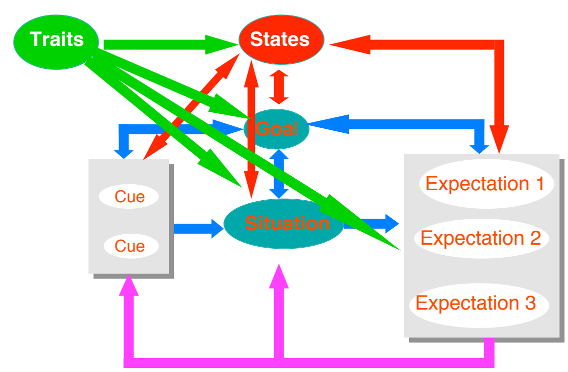

Figure 2-2: State and Trait Effects on Processing

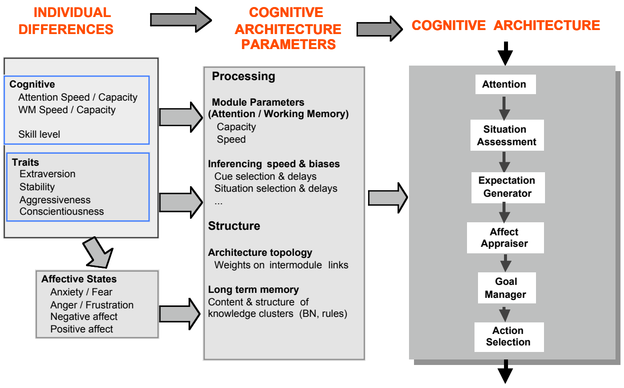

Figure 2-3: Parametric State / Trait Modeling Methodology

Affect Appraisal Process

The affect appraiser model (figure 2-4) incorporates elements of several recent appraisal theories: *multiple-levels* (Leventhal & Scherer, 1987; Sloman, 2003; Smith & Kirby, 2001) and *multiple stages* (Scherer, 2001), uses empirical data, and is embedded within a cognitive architecture, which controls an agent's behavior during a simulated task.

The *multi-level structure* generates both a *low-resolution assessment* of the current set of stimuli, in terms of a valence, and a *higher-resolution categorical assessment,* in terms of four of the basic emotions: anxiety/fear, anger, sadness, happiness. Its *multi-stage structure* uses both *universal elicitors* (e.g., novelty, threat level, pleasantness, unexpectedness), to generate the valence using an *automatic appraisal* (roughly corresponding to the largely 'hardwired', 'primitive' appraisal components), and more *cognitively-complex and idiosyncratic elicitors* (e.g., individual history, expectation- and goal-congruence), to generate a categorical assessment using an *expanded appraisal.* This latter process roughly corresponds to elements of the schematic and conceptual levels of Leventhal and Scherer. The automatic appraisal thus generates a valence (positive or negative), reflecting a generalized affective state. The expanded appraisal generates intensity levels for four of the basic[1] emotions: fear / anxiety [2]; anger, negative affect (sadness), and positive affect (joy, happiness).

The Affect Appraisal module consists of three stages: automatic appraisal, expanded appraisal, and current state modulator. *Automatic appraisal* emphasizes the stimulus properties to calculate the valence state (positive or negative); specifically, unexpectedness ("is situation part of current expectations"), novelty ("is situation part of individual history"), and the situation's intrinsic threat and "pleasantness" levels. Trait effects are included via a multiplicative factor, reflecting the agent's temperamental predisposition toward negative or positive states. For example, high extraversion and low neuroticism individuals tend to be predisposed towards positive affective states, whereas low extraversion and high neuroticism individuals tend towards negative affective states. The primary appraisal process therefore uses the values of these two traits to either reduce (in the former case) or increase (in the latter case) the valence value. In both cases, the modeler can modify the valence increment / decrement interactively, and observe its effect on processing.

Expanded appraisal emphasizes the influence of the agent's internal motivational context, by taking into consideration the congruence of the current situations and expectations with the agent's goals, the general level of success in achieving the current goals (e.g., number of goals met vs. failing), and the more idiosyncratic effects of the particular elicitors (e.g., previous experiences with a specific stimulus). This latter factor captures the domain-, task-, and individual-specific factors, by associating a set of specific eliciting situations and expectations with a particular emotion. These factors are encoded in a set of belief nets, associated with a particular emotion, for each agent. Each belief net calculates a number of values for a particular emotion, and these are then averaged to provide the value for that emotion for a given processing cycle. The belief nets vary in type and structure, reflecting the differences in individual histories (e.g., previous negative / positive experience in a particular situation), sensitivity to particular factors (e.g., considerations of own sense of competence or control in affect appraisal), and responsiveness (e.g., magnitude of particular affect generated in response to a set of situations or expectations). For example, a trait-anxious agent considers own competence and sense of control during appraisal to determine the level of anxiety.

Trait effects are included as above, with specific trait combination biasing towards a particular emotion (e.g., high neuroticism - low extraversion predisposes towards anxiety and negative affect (Matthews & Deary, 1998)). The current valence also influences the expanded appraisal, by contributing to the intensity of the valence-congruent emotions; thus negative valence increases the value of anxiety, anger and negative affect, while positive valence increases the value of positive affect. This stage of processing reflects the interaction between the automatic and expanded appraisal stages.

The expanded appraisal produces thus a vector of intensities for each of the four emotions. This representation, along with the multi-level appraisal structure, supports the representation of mixed, ambiguous, and possibly conflicting emotions, which are

---

[1] Some emotion researchers question the distinction between 'basic' and 'complex' emotions, but many emotion researchers identify a small set of basic emotions which include fear, anger, joy, sadness, disgust, and surprise.

[2] Most emotion researchers make a distinction between fear and anxiety, in terms of a variety of stimulus properties and behavioral characteristics. This model does not make these distinctions.

quite common in real life (Eid, 2001), but have not been adequately explored in models (Scherer, 2003). We are just beginning to explore the behavioral consequences of these complex representational possibilities.

In the current implementation, both appraisal types use linear functions, consisting of a weighted sum of the eliciting factors, to calculate the valence and the emotions. The contributions of the individual eliciting factors are controlled by their associated weights, which can be modified interactively by the model developer. This allows manipulation and tuning of the model, to reflect emerging empirical data and alternative theories regarding the mechanisms of appraisal and state / trait effects, as well as modeling of a wide range of individual differences (e.g., differences in sensitivity to the valence produced by the automatic appraisal can be explored by modifying the weight of the valence component in the expanded appraisal functions).

The *Current State Modulator,* the final stage of the appraisal process, consists of modulating the newly-derived valence and emotion values by the existing valence and emotions, generated in the previous execution cycle, thereby assuring smooth transitions among states. Traits exert an effect on this stage by influencing the ramp-up and decay rates of individual emotions, as well as the maximum intensities.

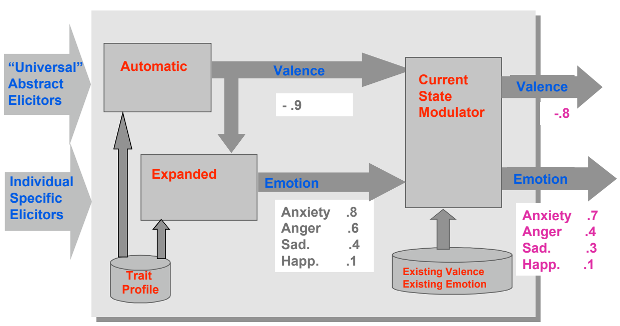

Figure 2-4: Affect Appraisal Model

The resulting affective states then influence processing in several ways. *First*, they are used directly in the rules selecting the agent's goals and actions. *Second*, they influence the speed and capacity of the architecture modules. *Third*, they influence mental construct ranking, thus determining whether a specific cue or situation is processed, or specific goal selected. The last two effects have been a particular focus of this effort, and aim to emulate some of the empirically-identified mechanisms of emotion effects within the perceptual and cognitive apparatus, as outlined above.

Results

The appraisal model described above, embedded within the MAMID architecture, was evaluated by modeling agent decision-making in the context of a peacekeeping scenario. The behavior of several commander types (anxious, aggressive, normal) was modeled by instances of the cognitive architecture. The simulated humans where exposed to a series of 'surprise' situations (e.g., destroyed bridge, hostile crowd), designed to elicit different reactions as a function of the state and trait differences. The agent's affective states were dynamically generated by the affect appraisal, in response to each surprise situation, and the results varied as a function of the agent's trait profile, as well as the current mental constructs representing the internal dynamic processing context (i.e., cues, situations, expectations, and goals).

The resulting emotions then influenced processing within each of the architecture modules, as outlined above (e.g., contributed to lower or higher processing capacities, threat bias, etc.). Figure 3-1 shows the fluctuating anxiety levels of a normal and a trait-anxious commander, during the course of the simulation scenario.

The different emotions, together with the trait-related differences in both processing and memory, then resulted in differences in behavior, in response to the identical set of external circumstances produced by the scenario simulation (figure 3-2), as outlined below.

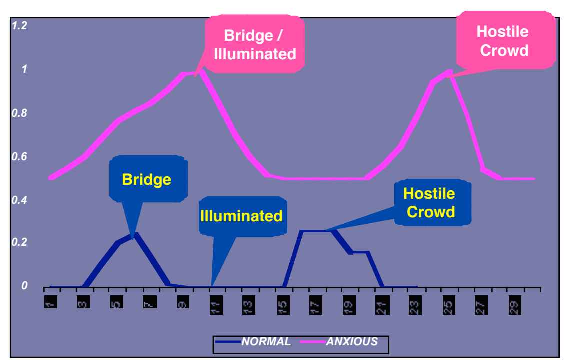

Figure 3-1: Anxiety Fluctuations Over Time for Normal and Anxious Commanders

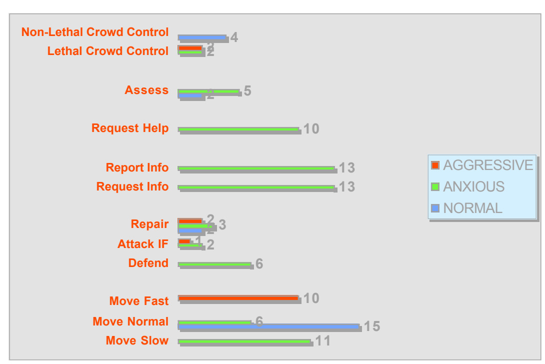

Figure 3-2: Distinct Behavior of Different Agent Types

Destroyed Bridge and Illumination

The first surprise event is a destroyed bridge, which must be repaired. This affords the first opportunity for differential reactions, depending on the agent's state and trait profile. The 'normal' commander conducts the necessary situation assessment, repairs the bridge, and moves on. In contrast, the 'anxious' commander stops, begins defending the unit (even though there is no direct evidence of the enemy), and engages in excessive assessment and communication, primarily as a coping strategy to reduce his heightened state of anxiety. Due to the consequent delay, he is still at the destroyed bridge when the second event occurs: enemy fires illumination rounds. (This event occurs at a fixed time frame of the scenario, frame 10, by which time the normal commander has already moved further along the route and does not react to this event, since it poses no immediate danger within the current environment.) The anxious commander now faces two anxiety-inducing situations: the destroyed bridge *and* the illumination rounds, which further increase his anxiety level. As a result of this high anxiety, he fires at the enemy, even though he has no knowledge of the enemy's location. He eventually does repair the bridge and moves on, but with significant delay and more slowly than his non-anxious counterpart.

Hostile Crowd

The next event encountered is a hostile crowd, which blocks the unit's route. The 'normal' commander realizes that this event does not represent imminent danger to the unit. It must, however, be dealt with so that the unit can proceed along the route and reach the objective. The commander therefore asks for assistance to disperse the crowd with peaceful means. In contrast, the 'high anxious' commander assumes that the crowd represents a real danger to the unit, and to himself, and overreacts by firing into the crowd. His primary focus is on reducing his own level of anxiety, by engaging in several 'coping strategies' (e.g., excessive communication and requests for help). These alternative behaviors are generated due to the distinct situations and expectations derived by the cognitive architectures representing the distinct commander types, which result in different affective states generated by the affect appraiser, and subsequent differences in activated goals and selected rules generating specific behavior.

The MAMID architecture provides a variety of alternative pathways to these outcomes, allowing for multiple sequences of mental constructs to result in the same final outcome, depending on the exact structure of LTM, the mental constructs derived in the current context, the affective states, and the module processing parameters. For example, the specific behavior of firing into the crowd may be triggered due to a specific goal to implement lethal crowd control, or to a rule which triggers this behavior in response to an 'under attack' situation and 'high anxiety'. This type of redundancy appears consistent with empirical evidence and allows the exploration of alternative mechanisms through which the effects of trait and state differences on decision-making and behavior can be expressed.

Discussion

These preliminary results indicate the feasibility of implementing a multi-level, staged appraisal process within a cognitive architecture, and demonstrating its ability to dynamically generate distinct emotions in response to distinct external and internal contexts. These emotions, in conjunction with distinct personality traits, then give rise to different behavior in response to the same set of external circumstances.

Can we thus say that we have constructed a model of human appraisal, and modeled the mechanisms of state and trait effects, albeit in a highly-abstracted and simplified manner?

Computational models are powerful tools in the emotion researcher's repertoire. It is possible to construct an agent model, with distinct emotions resulting from distinct external and internal circumstances. It would however be a leap of faith to then claim that such a model represents what actually takes place in the mind.

It is critical to distinguish between verification ("the simulation does what we programmed it to do") and validation ("the simulated model's structure, processing and output correspond to the natural phenomena we are attempting to model"). The results above fall largely in the category of verification. How then would we go about validating an appraisal model, what types of data would be required, and are such data available?

We can consider two types of generic validation procedures: input/output oriented (black box) and process-oriented. For the former, we would need to define the desired output for each type of input, for a particular component of the model. For example, we might define the desired emotions for a particular set of elicitors, and determine whether the model conforms to these expectations, using data such as those generated by Roseman (2001). This is the approach taken by Scherer with his GATE model (Wehrle & Scherer, 2001). We might test whether the attention capacity reductions due to anxiety correspond to empirical data on a particular task, and whether the human-attended cues correspond to those "attended" by the model. For a model embedded within an architecture, and performing a particular task, such validation would require targeted empirical studies within the specific task-domain to obtain the desired data.

Process-level validation is even more difficult. Here the objective is not just input-output congruence, but congruence with the actual processing mechanisms and structures that mediate the appraisal process, as well as the perceptual and decision-making processes. These types of data are notoriously difficult to obtain, particularly in humans. While methods exist for inferring some of these processes and structures (e.g., introspection, indirect assessments using priming, and, more recently, dynamic imaging such as fMRI and PET scans), it is difficult to unequivocally determine, for example, whether a given affective state triggered a specific expectation or goal, particularly in complex domains.

Problems Encountered

Three primary problems were encountered in this modeling effort. First, *lack of empirical data*: construction of computational models, particularly models of agent architectures in field settings (as opposed to simpler, isolated phenomena in laboratories) requires empirical data that are currently not available. Frequently we relied on extrapolations of existing data, and often encoded only qualitative effects. Target empirical studies are required to provide data for focused computational models.

Second, *effort required to construct LTM models:* the labor-intensive nature of human behavior model development is a recognized problem. Developing knowledge-acquisition tools, shared domain ontologies, and interactive GUI's helps somewhat, but does not solve the inherently difficult problem of attempting to emulate results of phylogenetic and ontogenetic development. Cognitive architectures capable of learning are likely to help somewhat.

Third, *model brittleness and the necessity for repeated fine-tuning* of both the knowledge-bases (agents' LTM), and the model parameters: this problem results from a combination of factors, including the use of symbolic representations, and the complexity of both the model and the task. As above, including learning would help somewhat, as might different representational and inferencing formalisms (e.g., connectionist or spreading activation techniques).

Conclusions and Future Work

The model described above has merely begun to 'scratch the surface' of appraisal modeling. Many extensions are possible, and necessary, to develop a more accurate model of the complex appraisal processes. These include the following: additional processing levels; parallelism and interaction among processing levels; more complex elicitors; inclusion of coping potential in elicitors (Lazarus' deliberate appraisal); more emotions (basic and complex); more complex functions calculating the architecture parameters and emotion dynamics; and explicit models of metacognition.

An important question however is whether this type of a model, embedded within a cognitive architecture and attempting to model human behavior in a complex domain, is an appropriate vehicle for exploring computational models of appraisal. One could argue that such models are more ecologically valid than models of 'disembodied' appraisal processes in laboratory tasks. But this advantage comes at a price: the effort to build the architecture, the task-specific long-term memory models, and the simulation environment, and the danger that during this process artifacts will be introduced which reflect the complexity of the modeling task itself, rather than the phenomenon being modeled.

Within the challenging area of appraisal modeling, both approaches have their place. What is needed is an improved understanding of the specific factors that call for one or the other approach, as well as a systematic exploration of their limitations and benefits.

Acknowledgements

This research was supported in part by US Army contract DASW01-00-C-3000. We would like to acknowledge support of the COTRs, Drs. B, Witmer and J. Psotka, and the contributions of: Dr. J. Pfautz, Profs. William Revelle and Gerald Matthews, and Mr. Ted Fichtl, as well as the Psychometrix software team.

References

Eid, M. 2001. Advanced Statistical Methods for the Study of Appraisal and Emotional Reaction. In *Appraisal Processes in Emotion*. K.R. Scherer, A. Schorr, T. Johnstone (Eds.). NY: Oxford University Press.

Ellsworth, P.C. and Scherer, K.R. 2003. Appraisal Processes in Emotion. In *Handbook of Affective Sciences*. R.J. Davidson, K.R. Scherer, H.H.Goldsmith (Eds.). NY: Oxford University Press.

Forgas, 2003. Affective Influencees on Attitudes and Judgments. In *Handbook of Affective Sciences.* R.J. Davidson, K.R. Scherer, H.H.Goldsmith (Eds.). NY: Oxford University Press.

Hudlicka, E. and Fellous, J-M. 1996. *Review of Computational Models of Emotion*. Technical Report 9612. Arlington, MA: Psychometrix Associates, Inc.

Hudlicka, E. 1998. *Modeling Emotion in Symbolic Cognitive Architectures.* AAAI Fall Symposium Series, TR FS-98-03. Menlo Park, CA: AAAI Press.

Hudlicka, E. 2002. This time with feeling: Integrated Model of Trait and State Effects on Cognition and Behavior. *Applied Artificial Intelligence,* 16 (7-8):1-31.

Isen, A.M. 1993. Positive Affect and Decision Making. In *Handbook of Emotion*s, J.M. Haviland, M. Lewis, (eds.). NY: Guilford.

Leventhal, H. and Scherer, K.R. 1987. The relationship of emotion to cognition. *Cognition and Emotion*, 1:3-28.

Matthews, G., and Deary, I.J. 1998. *Personality Traits.* Cambridge, UK: Cambridge University Press.

Roseman, I. 2001. A Model of Appraisal in the Emotion System. In *Appraisal Processes in Emotion*. K.R. Scherer, A. Schorr, T. Johnstone (Eds.). NY: Oxford University Press.

Scherer, K.R. 2001. Appraisal Considered as a Process of Multilevel Sequential Checking. In Appraisal Processes in

Emotion. In *Appraisal Processes in Emotion*. K.R. Scherer, A.Schorr, T. Johnstone (Eds.). NY: Oxford.
Scherer, K.R. 2003. Introduction: Cognitive Components of Emotion. In *Handbook of Affective Sciences.* R.J. Davidson, K.R. Scherer, H.H.Goldsmith (Eds.). NY: Oxford University Press.
Sloman, A. 2003. How many separately evolved emotional beasties live within us? In Emotions in Humans and Artifacts, R.Trappl, P. Petta, and S. Payr (Eds.). Cambridge, MA: The MIT Press.
Smith, C.A. and Kirby, L.D. 2001. Toward Delivering on the Promise of Appraisal Theory. In *Appraisal Processes in Emotion*. K.R. Scherer, A.Schorr, T. Johnstone (Eds.). NY: Oxford University Press.
Wehrle, T. & Scherer, K.R. 2001. Toward Computational Modeling of Appraisal Theories. In *Appraisal Processes in Emotion*. K.R. Scherer, A.Schorr, T. Johnstone (Eds.). NY: Oxford University Press.
Williams, J.M.G., Watts, F.N., MacLeod, C., and Mathews, A. 1997. *Cognitive Psychology and Emotional Disorders*. NY: John Wiley.