Appears in the Proceedings of the 23rd Annual Conference of the Cognitive Science Society, Edinburgh, Scotland

Modeling the Interplay of Emotions and Plans in Multi-Agent Simulations

Stacy Marsella (marsella@isi.edu)
USC Information Sciences Institute, 4676 Admiralty Way
Marina del Rey, CA 90292 USA

Jonathan Gratch (gratch@ict.usc.edu)
USC Institute for Creative Technologies, 13274 Fiji Way
Marina del Rey, CA 90292 USA

Abstract

The goal of this research is to create general computational models of the interplay between affect, cognition and behavior. These models are being designed to support characters that act in virtual environments, make decisions, but whose behavior also suggests an underlying emotional current. We attempt to capture both the cognitive and behavioral aspects of emotion, circumscribed to the role emotions play in the performance of concrete physical tasks. We address how emotions arise from an evaluation of the relationship between environmental events and an agent's plans and goals, as well as the impact of emotions on behavior, in particular the impact on the physical expressions of emotional state through suitable choice of gestures and body language. The approach is illustrated within a virtual reality training environment.

Introduction

Emotions play a central role in our lives. A wealth of empirical research has revealed a complex interplay between emotions, cognition and behavior. Emotional state may impact decision-making, actions, memory, attention, voluntary muscles, etc., which, conversely, may influence emotional state (e.g., see Berkowitz, 2000). Teasing apart and understanding these complex relationships is not an easy undertaking.

Not surprisingly, given this complexity, there are also a wealth of emotional models, with starkly differing views concerning the relation between cognition and emotion. While some theories have argued that cognition has a central role in evoking emotions (Lazarus, 1991), others have argued for a more minor role (Zajonc, 1984). With regards to the effects of emotions, theories of emotion have historically posited them as a problem for cognition, an impediment to effective cognitive function. On the other hand, more modern theories view emotions as more helpful than problematic, for example, a mechanism that facilitates human adaptation (e.g. Lazarus 1991, Simon, 1967).

We come to this conundrum from a certain perspective. The focus of our work is on general software agents that model human performance in rich simulated worlds. In particular, we focus on virtual training environments where intelligent agents interact with a human participant to facilitate the training objectives.

Emotions play an important role in such environments by enhancing believability and realism, increasing a sense of empathy and attachment to synthetic characters, and adding to the suspense of the simulation. For example, one of our environments, *Carmen's Bright IDEAS*, is designed to teach mothers of pediatric cancer patients better problem solving skills (Marsella et al., 2000). The mother learns by interacting with agents in a simulated world that mirrors her own. In particular, emotional models are used to help the mother identify with a human-like agent who faces various social problems due to her child's cancer. Another example is the Mission Rehearsal Exercise, a training environment designed to teach decision-making skills in highly evocative situations (Swartout, et al., 2001). The system provides an immersive learning environment where participants can experience the sights, sounds and circumstances they will encounter in real-world scenarios while performing mission-oriented training (Figure 1). Emotional models are used to enhance the intensity of the experience by creating characters that can respond emotionally to the student's decisions.

These simulations are set apart by the complexity of the environments and, more importantly, the detailed cognitive, emotional and behavioral modeling required. The agents face a variety of social and physical challenges, requiring the generation and execution of complex multi-agent plans. Overall, this complexity distinguishes this effort from more abstract simulation environments designed to study long term interactions of simpler agents (e.g., Nicholson et al., 1998) or believable, non-human agents in games (Neal Reilly, 1996).

Although complex, these realistic simulation environments offer a unique opportunity to explore and evaluate issues that arise by virtue of the complexity and fidelity of the modeling. For example, the agents must be able to generate complex plans with multiple goals and sub-goals. These plans may need to evolve or be replaced over time. Therefore, a key issue arises as to how the dynamics of this process and the structure of

Figure 1: A scene from the Mission Rehearsal Exercise

the resulting plan relate to overall emotional state and its dynamics. Another key issue concerns the agents' behaviors. They must interact with human participants across a range of modalities in a way that appropriately conveys their underlying emotional state. The wide repertoire of human nonverbal behaviors must be modeled, both subtle and extreme behaviors, consistent with emotional state. Fundamental questions arise as to what behaviors are exhibited and how various cognitive and emotional factors mediate between alternative behaviors. Finally, the realism of these simulations affords a unique, albeit weak, form of evaluation. The realism here supports more direct comparison with human behavior under matching conditions.

In essence, we are suggesting that it can be useful to attack the emotion conundrum head on via comprehensive, realistic simulations. Such simulations raise interesting research questions for cognitive science. Indeed the relation is synergistic since research on human cognition and emotion drives the design of our models.

In this paper, we demonstrate how some of the daunting subtlety in human behavior can be modeled by intelligent agents, from the perception of events in the world, to the appraisal of their emotional significance, through to their outward impact on agent behavior. We put forth a domain-independent solution that focuses on the problem of modeling "task-oriented" emotions – emotions that arise from performance of a concrete task. We then go on to illustrate the application of this model to virtual training environments.

Plans, Emotion & Behavior

The agents we design must provide convincing portrayals of humans facing difficult, dangerous problems. In particular, they must exhibit emotionally revealing nonverbal behaviors and expressions consistent with deeply evocative/disturbing situations. These behaviors must also change in concert with the emotional state of the agents; obviously people express themselves differently when sad, happy or angry.

Of course, one cannot realistically convey emotions without realistically modeling the genesis of those emotions. Because planning is central to our agent's behavior, we first needed to address how agents' plans/goals lead to their emotions. Then, we needed to address the impact of emotion on behavior. The driving force behind our modeling efforts was psychological research on the relation of cognition, emotion and behavior. However, the development of the models also raised significant research issues.

Plans and Emotional Appraisal

Many psychological theories of emotion emphasize the tight relationship between emotions and cognition. Emotions clearly influence our decision-making (Clore et al., 1994; Fiedler & Bless, 2000). What is less recognized is the strong influence cognition has over emotion. For example, the same event could evoke a variety of emotional responses depending on our mental state: getting a flat tire could evoke anger or joy depending on if we want to reach or avoid our destination. Such events derive their emotion charge, not from some intrinsic emotion evoking properties, but from our interpretation of their significance. Much of the recent theorizing on emotion builds on this observation, arguing that emotions arise from a cognitive appraisal of how events impact our plans and goals (Ortony et al, 1988; Lazarus, 1991).

Such psychological findings are problematic for building realistic models of human emotion. Just as fans of different teams will respond differently to the score of a goal, intelligent agents must respond differently to events in the simulation, and in a way that appears coherent to a human observer. For an agent developer, however, psychological findings and theories are seldom cast in a way that easily translates to general computational models.

Fortunately, there has been a nice convergence between cognitive appraisal models of emotion and the technologies underlying intelligent agents. Thus, while appraisal theories are vague on how events relate to goals, artificial intelligence planning methods now provide elaborate "mental" structures and inference techniques to assess this relationship (see Weld, 1999). While cognition cannot be reduced merely to planning, such algorithms can provide a cornerstone for making appraisal theories more concrete. By maintaining an explicit representation of an agent's plans, they can easily reason about future possible outcomes – a key requirement for handling emotions like hope and fear that involve future expectations. Planning techniques also detect interactions between plans, for example, as when the plans of one agent are incompatible with those of another – a key requirement for handling emotions like anger or reproach which typically involve multiple actors.

Modern planning techniques also support a rich model of how cognition influences one's emotional state. We can model some of the dynamic ebb and flow of human emotion by relating emotional appraisals to the current state of plans in an agent's memory. As plans grow and change through the planning process, so too the emotional state will change as a reflection of this process – in a sense providing a window into an agent's mental processes.

Finally, by providing an explicit and rich reasoning infrastructure, plan-based approaches facilitate models of how emotions impact decision-making. Emotional state can act as search control, focusing cognitive resources on specific goals or threats. It can also alter the overall character of problem solving. For example, negative emotions seem to lead to narrow focused problem solving while positive emotions lead to broader problem solving that attempts to achieve multiple goals simultaneously (Sloman, 1987).

Emotional State and Physical Behavior

Psychological research on emotion reveals its pervasive impact on physical behavior such as facial expressions, gaze and gestures (Argyle & Cook, 1976; Ekman & Friesen, 1969, 1971). These behaviors communicate considerable information about an individual's emotional state. This may be intentional, as in shaking a fist. On the other hand, behaviors such as rubbing one's thigh, averting gaze and raised eyebrows may have no explicitly intended role in communication, but they suggest considerable information about emotional arousal, attitudes and attention. Indeed, observers can reliably infer a person's emotions and attitudes from nonverbal behaviors (Ekman & Friesen, 1969). For example, depressed individuals may avert gaze and direct gestures inward towards their bodies. An angry person's nonverbal behavior tends, if unsuppressed, to align itself with the object of the anger (e.g., by confrontational stares or obvious avoidance of eye contact).

Such movements also serve to mediate the information available to the individual. For example, if a depressed individual's head is lowered, this also regulates the information available to the individual. Orienting on an object of fear or anger brings the object to the focus of perceptual mechanisms, which may have indirect influences on cognition and cognitive appraisal by influencing the content of working memory. Even a soothing behavior like rubbing an arm may serve to manage what a person attends to (Freedman, 1972).

These findings provide a wealth of data to inform agent design but such sources are descriptive, not proscriptive, often leaving open many details as to how alternative behaviors are mediated. Contemporary agent technology allows one to create rich physical bodies for intelligent characters with many degrees of physical movement. This forces one to directly confront the problem of emotional consistency. For example, an "emotionally depressed" agent might avert gaze, be inattentive, perhaps hug themselves. However, if in subsequent dialog the agent used strong communicative gestures such as beats (McNeill, 1992), then the behavior might not "read" correctly. Similarly, people don't tend to nonchalantly use deictic gesture while simultaneously averting their gaze due to mild feelings of anger or guilt. Such behavior may look un-natural, inconsistent, or may convey a different shade of meaning depending on context. Which is not to say that the overall mix of behaviors should always be monolithic. People do say one thing while expressing another. At the least, the mix of nonverbal behaviors often shade the meaning of what is said or communicated nonverbally. Returning to the previous example, if an agent does combine deictic gesture with gaze aversion, it may shade the interpretation dramatically, towards an expression of extreme emotion and a desire to control that emotion. For example, the agent is so disgusted with the "listener", they can't bear to look at them.

Implicit in these various concerns is that the agent has what amounts to a resource allocation problem. The agent has limited physical assets, e.g., two hands, one body, etc. At any point in time, the agent must allocate these assets according to a variety of demands, such as performing a task, communicating, or emotionally soothing themselves. For instance, the agent's dialog may be suggestive of a specific gesture for the agent's arms and hands while the emotional state is suggestive of another. The agent must mediate between these alternative demands in a fashion consistent with their goals and their emotional state.

Implementation

Implementations demand compromise. In our work we limit the scope of models by what agent technology currently does well, rather than trying to develop comprehensive but less general solutions. Thus, we focus on emotions arising from plan generation and execution, and ignore a number of potential sources of emotion, such as ego conflict. Similarly we focus on physical behavior, expressing emotion through body gestures and facial expressions, ignoring the myriad ways people communicate emotion through speech (and instead rely on pre-recorded voice clips for verbal communication).

An agent consists of three main components. The planner/executor maintains a representation of the world state, and develops, executes and repairs plans that achieve the agent's goals. STEVE (Rickel & Johnson, 1998) plays the role of the planner/executor in the application described below, but variety of AI planning methods could serve this role. The other components implement the cognitive appraisal of emotions and manage their physical manifestation.

Cognitive Appraisal

As we alluded above, we focus on cognitive appraisals as they relate to an agent's plans and draw on the strengths of modern artificial intelligence planning techniques. Specifically, we build on Émile, a computational realization of Ortony et al.'s cognitive appraisal theory (Gratch, 2000). The approach assesses the relationship between events and an agent's disposition (described by its goals, social standards). Unlike most computational accounts, Émile explicitly considers the role plans play in mediating the relationship between events and the agent's disposition. Rather than appraising events directly, Émile appraises the state of plans in memory, as inferred and elaborated by a general-purpose planning algorithm. This allows Émile to avoid the large number of domain-specific appraisal rules needed by prior computational approaches (e.g., Elliott, 1992). Domain-specific information, for the most part, is restricted to the operator descriptions (the domain theory) from which plans are built, and which an intelligent agent needs anyway to inform planning and action selection.

Émile also draws heavily on the explicit plan representation to derive the intensity of emotional response. Émile incorporates the view of Oatley and Johnson-Laird (1987) and Neal Reilly (1996) that emotions are related to changes in the perceived probability of goal attainment. Intensity is broken down into the probability of the event in question (e.g. the probability of goal achievement or the probability of a threat) and the importance (utility) of the event to the agent, both of which are derived from the current plan structure. As intensity is based on the current plans, the assessment is a reflection of their current state and changes with further planning. Individual assessments are aggregated into a set of "leaky buckets" associated with each emotion, where these buckets represent the current intensity of different emotions.

Physical Focus

The key challenge of the behavior component is to manage the flexibility in an agent's physical presence in a way that conveys a consistent emotional state. Agents are represented by rich bodies with fully articulated limbs, facial expressions, and sensory apparatus. The implementation must control the degrees of freedom provided by the agent's body in a way that satisfies the constraints imposed by psychological findings

To address this problem we rely on the Physical Focus model (Marsella et al. 2000), a computational technique inspired by work on nonverbal behavior in clinical settings (Freedman, 1972) and Lazarus's (1991) delineation of emotion-directed versus problem-directed coping strategies. The Physical Focus model bases an agent's physical behavior in terms of what the character attends to, how they relate to themselves and the world around them, specifically whether they are focusing on themselves and thereby withdrawing from the world or whether they are focusing on the world, engaging it.

The model organizes possible behaviors around a set of modes. Behaviors can be initiated via requests from the planner/executor or started spontaneously when the body is not otherwise engaged. At any point in time, the agent will be in a unique mode based on the current emotional state. This mode predisposes the agent to use particular nonverbal behavior in a particular fashion. Each behavior available to an agent is categorized according to which subset of these modes it is consistent with. Any specific nonverbal behavior, such as a particular nod of the head, may exist in more than one mode and conversely a type of behavior, such as head nods in general, may be realized differently in different modes. Transitions between modes are based on emotional state.

Modes also influence an agent's sensitivity to external stimuli, currently in a simplistic fashion. Rather than modeling the full flexibility of how people can focus their perception and attention (Wells & Mathews, 1994), we provide a domain specific mechanism for ranking stimuli by their intensity and filtering certain stimuli depending on if the focus mode is inner or outer directed.

Grouping behaviors into modes attempts to mediate competing demands on an agent's physical resources, especially gesturing and gaze, in a fashion consistent with emotional state. This grouping model is designed with the intent that it be general across agents. However, realism also requires that specific behaviors within each mode incorporate individual differences, as in human behavior. For example, we would not expect a mother's repertoire of gestures to be identical to that of an army sergeant.

In the current work, we model three modes of physical focus: body-focus, transitional and communicative (as opposed to the five modes discussed in Marsella et al., 2000). Body focus is marked by a self-focused attention, away from the conversation and the problem-solving behavior. Emotionally, it is associated with considerable depression or guilt. Physically, it is associated with the tendencies of gaze aversion, paused or inhibited verbal activity and hand to body stimulation that is either soothing (e.g., rhythmic stroking of forearm) or self-punitive (e.g., squeezing or scratching of forearm). The agent exhibits minimal communicative gestures such as deictic or beat gestures (McNeil 1992,

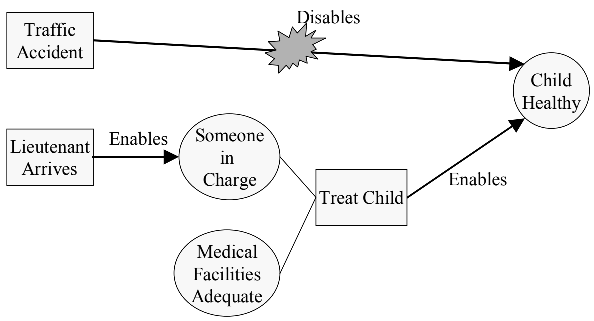

Figure 2: An example of the mother's plan

Cassell & Stone 1999) when in this mode. Transitional indicates an even less divided attention, less depression, a burgeoning willingness to take part in the conversation, milder conflicts with the problem solving and a closer relation to the listener. Physically, it is marked by hand to hand gestures (such as rubbing hands or hand fidgetiness) and hand to object gestures, such as playing with a pen. There are more communicative gestures in this mode but they are still muted or stilted. Finally, communicative indicates a full willingness to engage in the dialog and problem solving. Physically, it is marked by the agent's full range of communicative gestures, use of gaze in turn taking, etc.

Illustration

We illustrate the model by walking through an example of the system's behavior in the context of a virtual environment for familiarizing soldiers with the demands of peacekeeping operations. The Mission Rehearsal Exercise attempts to create an immersive learning environment through the integration of high-fidelity real-time graphics, intelligent agents, 3D audio and an interactive story whose outcome depends on the decisions and actions that participants take during the simulation.

In our working scenario, the system models a mix of three interactive and about forty pre-scripted virtual humans that play the parts of characters in the peacekeeping exercise. A human trainee commands a platoon of soldiers that have become involved in an automobile accident while driving to meet another platoon in need of reinforcement. The student must decide how best to allocate his forces between the conflicting goals of assisting an injured child and completing his mission, all under the watchful eyes of a "ZNN" cameraman.

Currently, only the character portraying the injured child's mother incorporates our emotional model. Figure 2 illustrates a simplified representation of the mother's plan at the opening scene in the scenario. The mother is waiting for the lieutenant (the student) to arrive, which she views as a precondition for her child to be treated. She is somewhat angry with the lieutenant, perceiving him as responsible for the accident (the domain-theory hard-codes an attribution that the lieutenant is responsible for "accident" task). This appraisal is moderated by the importance of the goal (high) and the likelihood of the threat cannot be overcome (moderate). Initially she believes the medical facilities are adequate to treat the child on scene, meaning she has the simple plan in memory that the lieutenant should arrive and her child will be treated, neither task being under her direct control. Since her child is hurt, a threat to an important goal, she has high levels of distress. The likelihood the treatment will be successful even if applied is relatively low (implying that there are many non-specific threats to its success) so she is also extremely anxious. The sense of hopelessness (and anxiety) leads her to have an inner-directed Physical Focus. Her body gestures are directed inward and she will not attend to most stimuli.

When the lieutenant arrives, the mother perceives that the sub-goal that someone is in charge is now attained and all non-specific threats associated with its attainment disappear. The probability that the child will be treated grows, and the mother's distress diminishes enough to transition her into transitional focus. Her gestures become more outward directed and she attends to more perceptual stimuli and her child.

Later in the scenario, the lieutenant orders one or two squads forward to reinforce the platoon downtown. The mother interprets this as disabling her sub-goal that the troops help her child. The strength of this interpretation is influenced by the number of squads the student orders forward (implemented by domain-specific rules that infer the probability of the disablement based on the number of moving units). The appraisal model treats this as a blameworthy event, causing the mother to become angrier at the troops. This anger is sufficient to transition her into communicative mode. The planner repairs the mother's current plan, deciding that imploring the troops to stay is a way of redirecting their behavior. Her body language in performing this action is colored by her body focus and anger level, either remaining seated and gesturing mildly or raising to a standing position and gesturing strongly (see Figure 1).

Discussion

This project is still in its early stages (the initial prototype was completed at the end of September 2000). From a research perspective the biggest limitation is the lack of evaluation. Is it a viable learning environment? Does the addition of emotional models increase the realism of the scenario? Do people find the character's reactions plausible? How do emotional models impact the learning experience? Our plan is to begin formal evaluations in the coming year in conjunction with other research groups in the psychology and communications departments at the University of Southern California. Our anecdotal feedback has been encouraging.

We have demonstrated the system to a number of military personal and those who served in Bosnia or Kosova seemed strongly affected by the experience. One U.S. Army Colonel began relating a related incident after seeing the demo, became quite emotional, and concluded by saying, "this system makes people feel, and we need that." In another anecdote, someone playing the role of the lieutenant became agitated when the mother character began yelling at him and when she wouldn't respond to his reassurances (she cannot be mollified when her anger exceeds some threshold).

Finally, there are a number of limitations in how the system infers emotional state that need adjustment or re-thinking in light of this application. As mentioned, cognitive appraisal only addresses emotions that arise from a concrete representation of plans of goals. We only weakly address the influence of emotion on perception and completely ignore the influence emotions hold over beliefs. Another key issue is the notion of responsibility. For example, whom should the mother blame for the accident? The troops? Herself? Our sense is she should have a shared sense of responsibility and that this sense should change dynamically, influenced by her emotional state and subsequent actions of the troops. Our treatment of anger is also too simplistic. Anger seems influenced by the extent to which we decide someone intended the offending action and the extent to which they show remorse or attempt to redress the offence. We suspect the explicit use of plans can assist in forming such assessments, but we are still sorting out how.

These limitations not withstanding, the integration of plan-based appraisal of emotional state with the Physical Focus model provides a great deal of architectural support for emotional modeling. Furthermore, anecdotal evidence suggests that people find the agent's emotions to be plausible, and, to our surprise, people occasionally responded emotionally to our agents.

Acknowledgements

This work was supported by the US Army Research Office under a grant to the Institute for Creative Technologies. The content of this article does not necessarily reflect the position or the policy of the US Government.

References

Argyle, M., & Cook, M. (1976) *Gaze and mutual gaze.* Cambridge University Press.

Berkowitz, L. (2000). *Causes and Consequences of Feelings*. Cambridge University Press..

Cassell, J. & Stone, M. (1999). Living Hand to Mouth: Psychological Theories about Speech and Gesture in Interactive Dialogue Systems. AAAI Fall Symposium on Narrative Intelligence.

Clore, G., Schwarz, N., & Conway, M. (1994). Cognitive causes and consequences of emotion. In Wyer & Srull (eds.), *Handbook of social cognition,* 2nd ed.

Ekman, P. and Friesen, W.V. (1971). Constants across cultures in the face and emotion. *Personality and Social Psychology*, 17(2): 124-129.

Ekman, P. and Friesen, W.V. (1969). The Repertoire of NonVerbal Behavior: Categories, Origins, Usage and Coding. *Semiotica* 1:49-97.

Elliott C. D. (1992). *The Affective Reasoner: A Process Model of Emotions in a Multi-agent System*. Ph.D Thesis (TR#32), Northwestern University.

Freedman, N. (1972). The analysis of movement behavior during clinical interview. In *Studies in Dyadic Communcation*, 153-175.

Fiedler, K. & Bless, H. (2000). The interface of affective and cognitive processes. In Frijda, Manstead & Bem (eds.), *Emotions and Beliefs.* Cambridge University Press.

Gratch, J., (2000). Emile: Marshalling passions in training and education. *Proc. of the 4th International Conference on Autonomous Agents* Barcelona, Spain.

Lazarus, R.S. (1991). *Emotion and Adaptation.* Oxford Press.

Marsella, S. Johnson, W.L. & LaBore, C. (2000). Interactive Pedagogical Drama. *Proceedings of the Fourth International Conference on Autonomous Agents.* Barcelona, Spain, 301-308.

McNeil, D. (1992). *Hand and Mind*. University of Chicago Press, Chicago IL..

Neal Reilly, W.S., (1996). *Believable Social and Emotional Agents*. Ph.D Thesis CMU-CS-96-138. Carnegie Melon University.

Nicholson, A.E.,, Zukerman, I. & Oliver, C.D. (1998). Towards a Society of Affect-driven *Agents. In Proceedings of the 20th Cognitive Science Society,* Madison, WI.

Oatley, K. & Johnson-Laird, P.N. (1987). Towards a Cognitive Theory of Emotions. *Cognition and Emotion*, 1 (1).

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

Rickel, J. & Johnson, L. (1998). Animated agents for procedural training in virtual reality: perception, cognition, and motor control. *Applied Artificial Intelligence (13),* 343-382.

Simon, H. A. (1967) Motivational and emotional control of cognition. *Psychological Review, 74*, 29-39.

Sloman, A. (1987). Motives, mechanisms and emotions. *Cognition and Emotion*, 1, pp 217-234.

Swartout, W., Hill, R., Gratch, J., Johnson, W.L., Kyriakakis, C., Labore, K., Lindheim, R., Marsella, D., Moore, B.,Morie, J., Rickel, J., Thiebaux, M., Tuch, L., Whitney, R. (2001). Towards the Holodeck: Integrating Graphics, Sound, Character and Story, in *Proceedings of the Fifth International Conference on Autonomous Agents,* Montreal, CANADA.

Weld, D. (1999). Recent Advances in AI Planning. *AI Magazine* 20(2): 93-123.

Wells, A., and Matthews, G. (1994). *Attention and emotion: a clinical perspective.* Lawrence Erlbaum.

Zajonc, R.B. (1984). On the primacy of affect. *American Psychologist.* Vol. 39, No. 2, pp 117-123.