Factors of Emotion and Affect in Designing Interactive Virtual Characters

Proceedings of the 2013 Edward F. Hayes Graduate Research Forum

Kellen Maicher

Department of Design, The Ohio State University

Introduction:

The term *affective* may be defined as anything that expresses or elicits *emotion*, which itself is the complex psychophysiological state of mind involving physiological arousal, expressive behaviors, and conscious experience (Myers, 2004). *Affect* (the experience of feeling emotion) is one of the three components of modern psychology that includes the *cognitive* (processing) and *conative* (instinctive) ("Affect", 2011). Emotions themselves are components of a larger categorical subset that includes affective concepts such as *mood, attitude, personality*, and the *motive* (Norman, 2004. pg 43). Cognitive scientist Marvin Lee Minsky suggests in his book *The Emotion Machine* that emotions are simply different ways of thinking and serve as unique problem-solving paradigms that allow us to learn more effectively ("The Emotion Machine", 2010).

There are many theories regarding the specific mechanisms and functions of affect and emotion, with research dating back at least as far as the philosophers of ancient Greece (William, 1884). Notable historical contributors include Rene Descartes (1641), Charles Darwin (1872), and more recently William James and Paul Ekman. Modern theories of emotion include a number of somatic, neurobiological, cognitive, and evolutionary models; most of which are fairly complex and overlapping. Although relevant and intriguing concepts, a comprehensive overview of these theories is beyond the scope and purpose of this paper. Instead, a focus on the implications and processes of emotion and affect as they relate to interactive character design will be explored.

The Emotional Influence:

As with theories on underlying cognitive processes, the influence emotion has in specific areas of research and application is also tremendously diverse. Given that emotions are an intrinsic part of the human condition, it's understandable that an emotional context can be applied to nearly every facet of human life and many fields of study. Examples of this influence can be found in even the most cursory of investigations.

In humanistic sciences such as psychiatry and psychology, emotions are studied for their role in various cognitive and mental processes. Affective research in education and instruction explores how emotions influence the process of learning. In healthcare, they are studied to determine their effects on healing and well-being. Researchers in the social sciences analyze emotions to understand the connections they form in various kinds of communication and social interaction. Other relevant areas include philosophy, economics, political science, music, and art (just to name a few.)

But even these examples constitute just a small fraction of the influence emotion has in the human experience. As famed author Dale Carnegie once wrote," *When dealing with people, remember you are not dealing with creatures of logic, but creatures of emotion*" (Carnegie, 1936).

With such a wide scope of application and influence it can be daunting to provide a comprehensive review of emotions and affect as they relate to virtual character design. Since interactive digital characters can theoretically be used in any of the aforementioned fields (and for a variety of purposes), an in-depth review of the potential applications and current research in all areas becomes somewhat impractical. Taking a more condensed approach, the thrust of this review focuses on identifying key concepts and current theories underlying the most relevant areas of affective research directly related to digital character design. That being said, such a review would be difficult without first discussing the fundamental role and process of emotion in the human experience.

Communicating with Emotions:

While there are many debates concerning the specific origin, nature, and purpose of emotion, it can be said with a fair degree of certainty that one of the primary functions is communication. This is especially evident when considering the array of mechanisms humans have for displaying, detecting, and interpreting emotion. Known as *affective display*, people outwardly convey their emotions as a means of expressing themselves. Furthermore, humans are "hard-wired" to read and interpret such expressions in others. Examples include interpreting posture, facial expressions, gesture, and verbal intonation.

This ability to emote and interpret emotions seems to be an inherent condition in humans, as many evolutionary biologists theorize that *Homo sapiens* evolved the ability to communicate emotionally in order to exploit certain adaptive advantages (Darwin, 1872) (Norman, 2004. pg 136). This communication provides a cognitive buffer that allows humans to deal with a continually changing environment, providing a coping mechanism that aids with decision-making and communication. While many specific emotional displays must be learned (children mimicking their parents laughter for example), the capacity to do so appears to be innate (Norman, 2004. pg 31).

Figure 1: A children mimicking the smile of its mother. Although the ability to smile is innately inherent in all humans, such expressions must be learned.

According to evolutionary emotional theory, prehistoric ancestors possessing the ability to communicate in this fashion were better able to engage in group-related tasks such as hunting, fighting, and gathering ("Emotion", 2011). They were also better able to communicate interpersonally, which assisted with functions such as learning and pair-bonding.

Some scientists postulate that the ability to communicate emotionally is critical to maintaining important bonds between teachers and learners (Cooper et al 2000). Communicating ideas, processing concepts and feelings, criticism, and questioning are all processes that require non-verbal interactive components (Knapp 1978).

However, this doesn't suggest that emotions only serve as a means of communication, as there is a great deal of evidence to indicate that feelings also play distinct roles in cognitive processing, decision making, and other internal processes (cf. Dittrich et al 1996, Picard 1997, Lisetti and Schiano 2000, Damásio 1994). Damásio (1994) demonstrated the importance of emotion in decision-making with studies of neurologically damaged patients with affected emotion systems. Although these individuals still had the ability to communicate with others, they had extreme difficulties making commonplace decisions as a result of their emotional deficiencies.

Studies such as these highlight the importance of emotions to internal cognitive processing. However, given that the topic of this paper focuses on affective interaction, a greater focus will be placed on affective display and the communicative nature of emotions rather than these internal processes.

The Affective Process:

Like all forms of communication, affective interaction involves communication and/or action between two or more agents. In this case, the term "agent" is used because an affective system doesn't necessary rely on humans for interaction. An artificial or partially artificial system can employ synthetic constructs such as non-player video game characters or even inanimate objects as simple as a child's toy. Although many natural affective systems typically take place between two or more human beings, the focus of this research centers on systems that incorporate synthetic components. As such, the term *agent* is used to generally describe one component of an affective system even though specific designations such as "user", "simulation" or "character" will be used on a case-by-case basis.

Emotional Connections:

With affective interactions, an *emotional connection* can be said to develop when one or more communicating entities has the ability to share, predict, or empathize with the other's emotional states. In its most simple form, this process typically involves the following steps:

1. **Agent 1:** Convey an emotion (affect display)
2. **Agent 2:** Perceive the emotion
3. **Agent 2:** Recognize the emotion
4. **Agent 2:** Relate to/ process the emotion
5. **Agent 2:** React/ response to the emotion

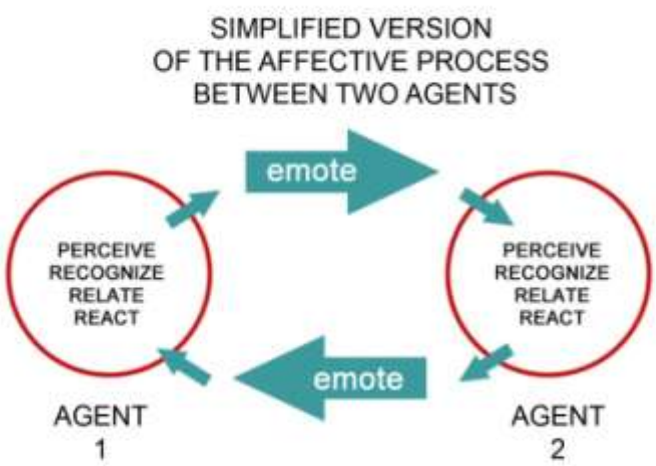

Figure 2: Simplified Version of the Affective Process

Of course, the process is much more complex than depicted in the figure above. Multiple signals and emotional cues are constantly being passed back and forth simultaneously, and each agent must attempt to perceive and process those signals into coherent messages (both instinctively and purposefully.) Although a powerful means of communication, the complexity of the affective process involves an incredible number of subtle nuances and unfortunately, errors in translation. With this in mind, it's understandable why many researchers in artificial intelligence and the cognitive sciences consider affect as an incredible challenge to understand, model, and implement.

Emotional Influences:

Generally, the emotions of agents can influence more than just a single additional entity ("Affect", 2011). While it's natural to think of communication taking place between two individuals, obviously there can be more people/ agents involved.

Figure 3: People arguing in public settings affect not only the mood of each other, but also the people in the surrounding environment.

Consider the example of a couple having a heated argument in a public space such as a restaurant. In such an environment there would almost certainly be other people in the room privy to the conversation (especially if it's of the heated variety.) Depending on the circumstances and relationships to the arguing couple, these bystanders will undoubtedly be affected by the emotional exchange in some way.

Additionally, it's important to understand that emotional exchanges don't simply influence emotions. Depending on the exchange, they can also affect additional characteristics such as mood, attitude, and behavior. Using the example of the arguing couple in a restaurant, bystanders may become upset or even amused at the emotional exchange being displayed. Their behavior may be altered in a loss of appetite or a desire to exit the restaurant more quickly than they normally would. Even long-term effects such as their general mood may be affected long after they leave the establishment.

This simple example emphasizes the impact emotional exchanges can have on the surrounding environment outside of the immediate point of exchange. In general, an agent's emotion can have an effect on four general areas ("Emotion", 2011):

- *Emotions of other persons/ agents*
- *Inferences of other persons/ agents*
- *Behaviors of other persons/ agents*
- *Interactions and relationships between the agent and other persons.*

Applying these concepts to the development of digital characters and synthetic agents, it's important to consider the conditions and environments in which these agents reside and function.

For example, in a massively multi-player online role playing game multiple users share the same virtual space. Therefore, emotional exchanges between two parties could have an effect on more than one player. In an era where networked interaction involving thousands and even millions of users is becoming more and more commonplace, such considerations must be taken into account by designers and developers.

An Emotional Frame of Reference:

For affect to take place, at least one of the agents in an emotional exchange must be able to a) accurately perceive and recognize the emotions b) process the emotions cognitively and c) relate to those emotions to some degree. This last component may be overlooked by designers, even though it is vital to the process of affective interaction.

In any affective system, the receiving agent must share a common frame of reference of the emotion(s) being displayed. This frame of reference is typically provided by the human user, although research in emotional modeling and artificial intelligence is quickly making progress in the creation of artificial systems capable of making emotional references to information stored in databases ("Artificial Intelligence", 2011).

To highlight this concept of "frame-of reference", consider the simple example of a human user interacting with an artificial video game character designed to express the emotion of fear in a virtual environment. The human user must not only recognize the emotion of fear through various cues such as facial expression, body posture, gestures, and speech, but they also must have context for the emotion of fear itself. Since basic fear is a universal concept among humans, generally this isn't a particularly significant problem. However, if the virtual character is expressing fear of exotic stimuli not typically found in real-world scenarios (say aliens or giant robots); the likelihood that the user would relate to the fear expressed by the digital character is lessened.

If this occurs, the opportunity for greater emotional connection decreases. By and large, the more contextual reference a person has to the emotion in question, the greater the chance for the user to relate to that emotion. Although this may seem like a fairly obvious assertion, this author found limited evidence to suggest that emotional contextual reference has historically been a major concern for character designers (although further investigation on the part of the author may be needed.)

Natural vs. Artificial Emotions:

Of course, emotions conveyed in an affective system aren't required to be "real" in the sense that they're generated or detected by humans. Obviously, synthetic agents used in artificial systems such as simulations and video games don't generate "real" human emotions, they approximate them. This begs the question of whether or not human users can accurately process and interpret synthetic emotions.

While humans likely evolved to interpret emotions from other humans, the complexity of the human mind allows us to expand beyond that particular limitation. Humans naturally have the ability to assign emotions to a variety of non-living objects, a process known as *animism* (Piaget, 1933). Although there are several theories on why humans engage this process, some researchers suggest that it allows people to better cope with objects and scenarios they don't fully understand (Reeves and Nass, 1996). By assigning emotions and other human attributes to non-human objects, we create a path toward understanding those objects by assigning attributes such as "purpose" and "intention" to explain as-yet unknown concepts in function and mechanism.

Figure 4: Avid gardeners can grow very attached to their favorite plants.

For instance, take the example of a person's emotional attachments to their favorite house plants. These plants are neither sentient nor immediately responsive to stimuli, yet some people assign them humanistic qualities. They may name them or even speak to them just as they would a pet or even a child. They feed and water the plants by thinking in terms of what the plant "wants" or "desires", even though these are concepts inherent only to sentient beings. When the plant bends or grows towards sunlight, they may subconsciously think that the plant is deliberately moving towards something it wants rather than simply undergoing a biomechanical process.

This isn't to suggest that there are droves of people who believe plants are sentient. Rather, the point here is to highlight the idea that assignment of emotions is the first natural step humans take when attempting to understand something.

Anthropomorphism is a related concept, which describes our tendency to assign physical attributes to inanimate objects. A classic example of this phenomenon is our ability to see faces in common, everyday objects. In actuality, people aren't seeing faces at all, yet they perceive them nonetheless.

Figure 5: Faces, Faces, Everywhere – Human instinctively see faces where none actually exist.

This process is especially evident when we see how children view the world and the objects within it. As a point of fact, they are particularly adept at perceiving life (Scassellati, 2000). As they do, they frequently assign life-like qualities to things that we clearly distinguish as inanimate, and we consider these actions natural because we understand that this is the process by which children learn.

As children age and obtain a more complex understanding of the world around them, this process is tempered by knowledge and experience. But as humans continue to speak to their pets, name their vehicles, and carefully preserve our childhood toys throughout adulthood, it's clear that this behavior remains partly with all of us for the rest of our lives.

At this point the concepts of affective, emotion, anthropomorphism, animism start to become blurred. Even so, it's evident that the ability to create emotional attachments to non-human objects allows us to explore affective paradigms in a variety of scenarios and for a multitude of applications. This innate human ability has had a profound and direct effect on research in a number of areas including interface design, robotics, and interactive character development (Breazeal 2003, Picard 1999, Poggi and Pelachaud 2000).

Affective Computing:

Affective computing may be defined as *"the study and development of systems and devices that can recognize, interpret, process, and simulate human affects."* While it could be said that this area began thousands of years ago, the modern impetus for this field started with Roseland W. Picard and her 1995 paper entitled *Affective Computing* (Picard, 1997).

Affective computing research primarily involves the three components of a) displaying emotion b) recognizing emotion and c) processing emotion. While the first component is still a major factor of affective computing research, it seems most recent interest appears to revolve around the latter two areas.

A possible explanation for this recent lack of research interest in the area of machine emotions and affective display may reside in the fact that animators, artists, and roboticists have already been working on emotive expression for some time. Animators have been creating expressive characters since the first part of the 20th century, and roboticists have been striving to create emotive automatons for decades. Likely, this results from the fact that the illusion of emotional expression is much easier to create than recognition and processing of emotions can be. When Mickey Mouse expresses happiness, it's not in response to some external stimuli like Pluto rushing to greet him at the doorstep. Disney animators create the appearance of spontaneously generated happiness, but not true happiness as humans experience it.

Affective Computing Research Areas:

Technical areas of applied affective computing research seem to focus on four basic areas:

- **Speech**
- **Facial expression**
- **Gesture**
- **Form/ Aesthetics**

Speech is a very strong positive affective trigger (Persson et al. 2000). Most speech research tends to center on recognizing and processing the emotions of users by recognizing their speech patterns. Important factors such as speech rate and pitch variables are analyzed through various methods of pattern recognition (Dellaert, et al, 1996) (Lee, C.M, et al, 2001).

Speech is only one of the communication channels at the disposal of those designing digital character interfaces. When humans interact, we continually observe and interpret each other's visual cues (Strongman, 1996) and each day we use hundreds of expressive movements

(Morris et al, 1979). In fact, gestures and other forms of non-verbal communication function to reinforce, replace, and control speech (Knapp 1978).

Facial expression is an extremely important emotional communication channel, and some researchers would contend it's the most important (Ekman and Friesen 1975). Unlike other channels, facial expression provides feedback, highlights mood and attitude, and next to speech is generally considered the most critical source of communication. Given these attributes, it's understandable that humans pay particular attention to the face when communicating (Knapp 1978).

Facial expression detection and analysis research uses various processes such as optical flow, hidden Markov modeling, neural network processing, or active appearance modeling ("Affective Computing", 2011). Gesture detection research takes many forms, from overall body gesture analysis to more focused areas of the body such as hand and limb-based gestures.

Affective Design:

A concept that draws from the field of affecting computing and human-computer interaction, *affective design* focuses on the idea of creating emotional relationships between users and products via the physical attributes of those products. In his book *Emotional Design: Why We Love (or Hate) Everyday Things,* Donald Norman attempts to encapsulate the field of affective design, its attributes, and fields of application (Norman, 2004).

Norman begins with his observation that functionality and usability historically have been deemed irrelevant with respect to aesthetics and emotion. The next section of the text is spent illustrating the critical nature of emotion to the field of design, beginning with the assertion that the mind is comprised of three functional levels that all designers must aspire to be successful with (Norman, 2004. Pg 21). The first level consists of the *visceral* – the automatic behaviors that occur in the pre-conscious state. This is the stage where appearances of objects and first-impressions matter most, and in this paper's context would most directly relate to the outward form of digital characters. By his own admission, Norman borrows this concept from Picard's paper *Affective Computing* (Picard, 1997), although he expands on this concept with his own variation of the concept.

Since this level works independently of cognitive thought, it instead operates on what Norman refers to as "*pattern matching*", a bottom-up process that relies more on instinct than recall or experience. He lists several examples of conditions that provide patterns for positive affect including: *warm, well-lit places, temperate climates, sweet tastes and smells, bright, highly saturated hues, soothing sounds, smiling faces, rhythmic beats, attractive people, symmetrical objects, and rounded, smooth objects.* These examples may produce negative automatic affects include *heights, sudden/ loud sounds, bright lights, looming objects, crowded dense terrain, crowds of people, and sharp objects.*

It's important to note that these examples consist of Norman's personal opinion and not tested stimuli. Even so, these examples and the visceral levels they depict may be useful considerations in regards to affective character design. By applying targeted and compelling visceral attributes to our digital characters such as distinctive colors, appropriate sounds, and compelling voices designers may be more likely to create characters that are successful in this "first impression" stage. This may be why video game characters (the games of which are primarily targeted at young males that highly influenced by visceral phenomena), instinctively employ so many of these attributes.

The next level of the mind consists of the *behavioral* and includes everyday behaviors involving experience with and use of a thing. Specifically, this level relates to function, performance, and usability. Given interactive digital agents such as game avatars, this level would correspond to functional attributes such as how well characters move and respond to user input. Given this consideration, the game designer should employ all the concepts of functionality and usability at their disposal to make game characters work as they are intended to.

Finally, the last level in Norman's taxonomy is the *reflective* level of the mind. This is where most emotional and cognitive processes take place. In fact, it's the most relevant to affective theory since emotional response is dependent on processing and interpreting other stimuli and emotions. It's also the level most vulnerable to variability in terms of concepts like culture, experience, and education (Norman, 2004. pg 38) and one that has the ability to override the other two levels. For example, while most people may not be particularly fond of insects and spiders (a visceral response), entomologists employ reflection and their previous experience and knowledge to develop a fascination and appreciation for such creatures.

Norman goes on to state that successfully integrating all three levels in affective design is not an exact science and difficult to practicality implement. While he describes these three levels as universal concepts, he readily admits that human variability and distinctions between individuals makes it difficult to employ a "one-size fits all" approach to affective design. Depending on the particular application, it's likely a give-and-take approach must be employed when considering these levels for affective character design.

In the second third of his text, Norman discusses the consideration of these levels to the following areas he considers important to the field of affective design:

- **Personalization**
- **Customization**
- **Appropriateness of setting for objects**
- **Evaluating and distinguishing between user wants and user needs**
- **The importance of objects that evoke memories**
- **Establishing feelings of self**
- **The importance of product personality**
- **Designing for fun**
- **The importance of music and other sounds**
- **Establishing user trust with products**

Affective Virtual Characters (AVCs):

For the purpose of this review, *Affective Virtual Characters (AVCs)* may be defined as representative digital interfaces that possess the affective characteristics of emotional display, perception, recognition, and processing. While there are many forms of affective systems, the word “character” is important in this context because it defines a system using characters similar to those found in traditional media such as film or animation. This distinction is relevant because such characters typically have human-like qualities like personality and emotion. Thus AVCs are characters that function as interfaces but resemble living beings visually, cognitively and emotionally.

Affective virtual characters can be distinguished from other affective systems in that they serve as digital representations of real-world people or objects. For example, a text-based chatter-bot would not qualify using this criterion because such an application lacks a representative visual/ graphic component. The use of representative interfaces is critical to the concept of AVC's because of the advantages they convey with familiarity, interaction, and recognition.

Figure 7: Web pages and applications may one day be embedded with affective responses, but the lack of a representative interface may make it difficult to relate to and communicate with.

The line becomes somewhat blurred when referring to affective systems that have graphic interfaces that aren't necessarily representative of real-world objects. For example, typical web applications provide a graphic-user-interface (GUI) based on non-representative forms such as buttons, links, and menus (see Figure 7). Even if interfaces such as these were imbued with affective characteristics such as emotional responses, it would be a challenge to categorize them as AVCs because they lack a representative graphic interface such as a human avatar.

The advantage of using representative forms such as digital characters for interaction is that humans innately understand how to communicate with such forms, especially emotionally. When a digital human raises its eyebrow with an expression of surprise, we instinctively understand what this expression means. A typical Web GUI such as the one shown in Figure 7 would have to provide some other means of conveying the emotion of "surprise." Examples of possible alternatives could include using text or audio forms of communication such as a pop-up window with the text "I wasn't expecting that!"

Still, the distinction between affective systems and AVCs can be difficult to make. For example would an automated text-based customer service representative like the one shown in Figure 8 qualify? Assuming it had the ability to detect and respond to user emotions, would the fact that it uses text rather than speech disqualify it from this category?

Figure 8: Does an application that uses text instead of speech separate it from an ARC?

Robots would almost certainly meet the criteria as they are just as capable of affective processes as digital characters are. However, since the focus of this review centers on virtual characters, robots will be left in a class by themselves (perhaps a*ffective robotic characters* or *ARCs*.)

Embodied Conversational Agents (ECAs) are another category of agent that could conceivably fall under the umbrella of affective virtual characters. Defined as agents with representative graphic interfaces capable of engaging in conversation with humans and other agents using the same methods as humans do, these constructs very nearly fit the definition of an AVC.

However, two important features may set the two categories apart. First, the review of literature found no mention of emotion or affect being embodied in ECAs. While they are certainly defined and driven by intelligence, it seems affect is not a primary consideration. Whether or not intelligence implies affect and emotional content has yet to be determined, but it seems as though an agent capable of carrying on a conversation effectively should possess some degree of affective qualities.

The second characteristic that may separate ECAs from AVCs is the fact that it seems they are only intended to serve as autonomous agents. On the other hand, AVCs can consist of both user-governed agents (ex. avatars) and independent synthetic agents (ex. non-player video game characters.) As such, embodied conversation agents could be categorized as a subset of AVCs (assuming they are capable of affect and emotion.)

In general, a variety of constructs could conceivably fall under the umbrella of affective virtual characters depending on the context for which they are used. These could include video game characters, virtual avatars, and various anthropomorphic interfaces. While the specific designation and use may vary, the general properties of a) having a representative visual interface defined by a character and b) affective qualities are the properties that distinguish affective virtual characters from other affective systems.

Robotics:

Research efforts to create robots that mimic human expression, emotion, and response have existed long before computer graphics were sophisticated enough to allow us to create digital characters. As such, there may be insights to be drawn from in the field of robotics that could be applied to the creation of affective virtual characters.

Figure 9: Honda's ASIMO

As previously mentioned, certain types of robots would almost certainly qualify as affective characters. They are every bit as capable of affective interaction, and perhaps even more so because they are physical constructs capable of interacting with the real world. The fact that they must pass the "acid-test" of real-world interaction may lead to insights in form, movement, and expression that can be translated to digital characters.

The condition of functioning in a real-world environment offers several challenges to roboticists, not the least of which is the mechanical functions they must accomplish. Such constructs must overcome the problems of the physical world that digital characters do not have to contend with such as gravity, wind, friction, and mass. Many devices must employ sophisticated proximity detection and motion sensing mechanisms. Additionally, they must have analogous components to facial expressions, gestures, and body posture even if they don't possess humanoid features. In effect, robots must contend with all of the challenges virtual characters do as well as the considerable problems of functioning in the real world.

In spite of these challenges, the ability to exist in real environments offers several advantages to operating solely in a virtual setting as digital characters do. These advantages include opportunities for haptic feedback and real-world functionality that digital characters cannot provide.

Like digital characters, the process of creating affective robots is still in its infancy. It has been theorized that robot emotions may likely evolve from programmed survivability functions (Norman, 2004). The same programming that robots use to maintain balance, avoid falling off ledges, or running into walls may one day become sophisticated enough as to resemble the emotions of fear or apprehension. Code that allows them to be attentive and responsive to human users commands could evolve into something resembling loyalty or affection

Naturally, these types of "baby-step" approaches to programming affective responses could just as easily be employed in affective virtual characters. But the fact that robots must be able to function affectively in real space may provide insights to affect that are difficult to explore in virtual environments. By observing how robots move and respond to humans and the physical world around them, we may better understand how to make digital characters more effective.

Advantages of Affective Virtual Characters:

AVCs provide the opportunity to take advantage of two important human communicative traits: 1) our instinctive ability to relate to living things and 2) our natural tendency to assign life-like characteristics to non-human entities. In a sense, affective virtual characters bridge the gap between objects we've naturally evolved to communicate with (humans, animals, etc.) and synthetic machines that we haven't.

While digital characters have existed for decades, only recently have we begun to develop the ability to imbue affect and emotion in such creations. While traditional components of digital characters such as compelling visual aesthetics and expressive motion are still important qualities, these attributes will only take these characters so far. By instilling characters with affective traits, it may be possible to increase our ability to relate and communicate with these constructs.

Depending on the specific type of AVC in question, some or all of the following advantages may apply when using these interfaces.

- **Conversational turn-taking (with text or speech-based input)**
- **Facial expression of emotions**
- **Increased believability and trustworthiness of synthetic agents**
- **Information structure and emphasis**
- **Visualization and iconic gestures**

- **Social interaction and social conventions**
- **Orientation in a three-dimensional environment**
- **Increased user engagement and entertainment**
- **Verbal and non-verbal channels such as gaze, gesture, spoken intonation and body posture.**
- **Increased perceived usefulness of tasks and usability**
- **Non-distracting face-to-face communication that can be conducted in conjunction with other tasks**
- **Improved recall of information presented**

Additionally, some research indicates that users prefer visual cues of character status or emotion as opposed to verbal or text-based alternatives ("Embodied Agent", 2011). For example, a digital character expressing frustration is more likely to be satisfying to a user than a text-based message that says "I didn't understand the question. Please try again." Obviously, such visual cues are easier to achieve with representative interfaces such as those found in AVCs.

AVC Expression and Affective Display:

Due to their natural and simultaneously complex graphic representations, affective virtual characters have the ability to communicate and emote in a variety of ways that are instinctive to human users. These consist of one or more of the following:

- **Form**
- **General movement** – of the character as a whole
- **Component movement** (such as arms, legs, head, etc.)
- **Posture**

- **Gesture**
- **Speech/ Sound**
- **Facial Expressions**
- **Autonomous Behavior** – ticks, swaying, tapping, finger drumming, etc…
- **Reactionary Behavior** – in response to the user

Nathanson (1992) indicates that there are nine basic affects that influence people, each having biological expressions that accompany them:

Positive:

- **Enjoyment/Joy -** smiling, lips wide and out
- **Interest/Excitement** - eyebrows down, eyes tracking, eyes looking, closer listening

Neutral:

- **Surprise/Startle** - eyebrows up, eyes blinking

Negative:

- **Anger/Rage** - frowning, a clenched jaw, a red face
- **Disgust -** the lower lip raised and protruded, head forward and down
- **Dissmell** (reaction to bad smell) - upper lip raised, head pulled back
- **Distress/Anguish** - crying, rhythmic sobbing, arched eyebrows, mouth lowered
- **Fear/Terror -** a frozen stare, a pale face, coldness, sweat, erect hair
- **Shame/Humiliation** - eyes lowered, the head down and averted, blushing

Argyle (1988) asserts that non-verbal communication takes place by means of *posture, bodily contact, gaze and pupil dilation, general appearance, spatial behavior, clothing, or non-verbal vocalization.* Although attributes such as bodily contact may be inapplicable for digital characters, much of the rest of these points appear relevant.

B.J. Fogg proposes in his book *Persuasive Technology* that people use five social cues to infer sociability which could be used as a means of design affective characters (Fogg, 2003). These cues include:

1. **Physical:** eye, body, movement, etc…
2. **Psychological:** preferences, humor, personality, feelings, empathy, etc..
3. **Language:**
4. **Social Dynamics:** turn-taking, cooperation, praise for good work, answering questions, reciprocity
5. **Social Roles:** doctor, lawyer, Indian chief, etc…

Michael Schmitz of the German Research Center for Artificial Intelligence suggests that the criteria for creating <u>any</u> lifelike objects should include the following attributes (Schmitz, 2011):

- **Visual Appearance** aesthetics and first impressions
- **Presence of Voice**
- **Physiological Sounds -** sighs, grunts, groans, etc…
- **Pro-Active and Autonomous Behavior**
- **Stereotypical Behavior -** shyness, curiosity, etc…
- **Social Dialogue** - small talk
- **Social Deixis** - social protocol

- **Empathy**
- **Continuity Behavior**
- **Humor**
- **Needs and Wants**
- **Personality**

Pro-active and autonomous behavior refers to actions unprompted by the user, and are attributes supported by the work of Persson et al (2002) and White (1995.)

Schmitz notes that presence of voice should be used appropriately, as people tend to be uncomfortable hearing speech from simple physical objects that aren't supposed to speak such as lamps or shoes. Based on work by Wasinger et al., the idea is that users are comfortable talking with objects perceived to be complex such as digital cameras and computers, but simple objects feel unnatural (Wasinger, 2005.) However, given advancements in ubiquitous, ambient, and embedded computing it may be likely that people will become more accustomed to "smart objects" in the near future.

Schmidt also asserts that zoomorphic attributes such as basic needs and desires can be integrated into constructs to maximize life-like effects. He cites such examples as *preservation, pain, hunger, desire for sleep, exhaustion, fear,* and *excitement.* He again references the work of Persson et al. (2002), reinforcing the idea that adding such characteristics plays to basic human psychology that innately responds to animalistic behavior.

Facial Expression:

As previously stated, facial expression is particularly important for emotional communication. Many models and methods for facial animation are based on muscle systems, mimicking the structure and function of these organs. As such, it's important to understand how the face functions in displaying emotions. Muscles of the face are usually organized into five groups (Pioggia et al 2002):

- **Scalp and eyebrows** - Epicranius occipitalis, Epicranius frontalis
- **Eyes and eyelids** - Levator palpebrae superioris, Orbicularis oculi, Corrugator
- **Lips and mouth** - Quadratus labii superioris, Quadratus labii inferioris, Caninus, Triangularis, Zygomaticus, Buccinator, Mentalis, Orbicularis oris, Risorius
- **Nose -** Procerus, Depressor septi, Nasalis, Dilatator naris posterior/anterior.
- **Outer ear** - Auricularis anterior, Auricularis superior, Auricularis posterior

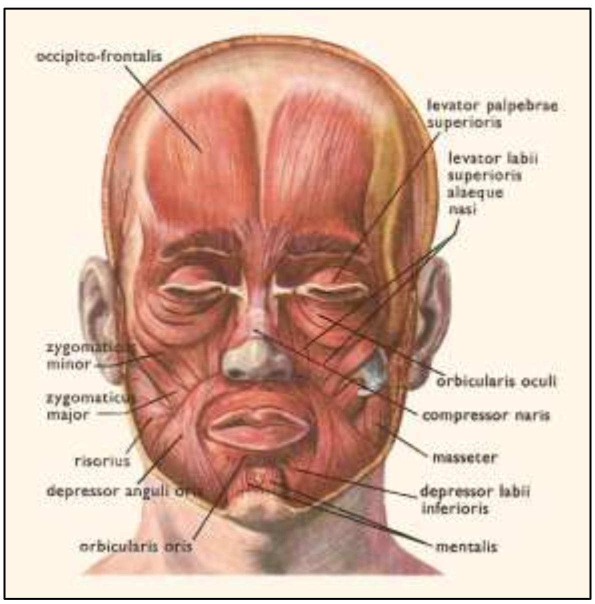

Figure 10: Muscles of the face. (Source: http://www.daviddarling.info/encyclopedia/S/skeletal_muscle_groups.html)

Taking advantage of grouping such as these, the *Facial Action Coding System (FACS)* was designed by Paul Ekman and Wallace V. Friesen in 1978 to label and quantify human facial expressions. The system was developed using a combination of surface *physiognomy* (features, movements, etc.) and sub-surface *physiology* (muscles, bones, etc.) FACS has been used by many studies in areas including computer generated imagery, facial animation, and various subsets of the social sciences (cf. Argyle 1994, Bartlett 1998, Ekman and Rosenzweig 1998, Terzopolous and Waters 1993, cf. Waters 1987, Yacoob and Davis 1994).

With FACS, programmers can manually code almost every possible facial expression, deconstructing it into the specific Action Units (AU) of which the system is comprised. As the FACS manual is over 500 pages in length, a detailed analysis is not appropriate for the purpose of this review. (However, the author intends to examine the FACS system in detail for potential integration into future applications.)

Ekman et al (1972) determined that there are six universal facial expressions:

- **Surprise**
- **Anger**
- **Fear**
- **Happiness**
- **Disgust/Contempt**
- **Sadness**

While these expressions vary in intensity and timing in various cultures (Zebrowitz 1997), they have been shown to be fairly universal as they can be found in all populations of the world (Argyle 1994). In terms of their variability, cultural norms known as **display rules** govern the degree to which individuals in a particular culture manage these expressions.

Historical approaches to modeling and animating the human face generally fall into two categories; **parameter-based** models and **muscle-based** models. Parameter-based models focus on controlling appearance while muscle-based models center on mimicking human muscle and skin.

The advantages of parameter-based models lie in the fact that they are not computationally-intensive but still offer a wide range of expressions. Hence, they are frequently used in many real-time applications such as video games and conversational agents. Muscle-based models offer realism and accuracy at the cost of being computationally intensive and time-consuming to implement.

Realism vs. Stylization and the Uncanny Valley:

Digital characters can be designed in all manner of shapes and sizes. They can be highly realistic in terms of form and function like *Benjamin Button* (Figure 11) or they can be stylized representations such as Pixar's *Mr. Incredible* (Figure 12).

Figure 11: The CG version of Benjamin Button from the film The Curious Case of Benjamin Button

When people think of designing complex affective virtual characters, they first may be inclined to pursue more realistic representations. Since the one of the goals of creating affective systems is to mimic real emotions and responses, it's natural to assume that the digital character itself must be as "real" as possible.

Figure 12: Pixar's Mr. Incredible

However, in reality there's evidence to support the idea that attempting to produce realistic characters is unnecessary (Benford et al, 1995). Some researchers suggest that it's more important to develop natural behavior rather than realistic form and appearance (Bailenson, 2005). A fairly recent study on the effect of copresence and avatar realism (Garau et al, 2005) demonstrated no measurable advantage to more photorealistic avatars.

Still, the matter of realism seems to be a point of contention. Discussions in this regard inevitably bring up the *Uncanny Valley* effect, an as-yet unproven theory that states that the closer human constructs approach realistic human form and movement, an innate feeling of unease, rejection, and even revulsion may occur. The supposed rationale for this reaction is the observation that people are so adept at spotting imperfections in the human form that they are repulsed by constructs that come close to a perfect reproduction (but aren't 100% successful.)

There is some scientific evidence to support the Uncanny Valley theory. A study by MacDorman and Ishiguro (2006) observed user reactions to images of facial morphing from a mechanical robot - to a more human robot - to a real human. They determined that users reached the peak of greatest unease where the figure appeared slightly less than human. However, some critics argued that other factors may have played into these results, including the attractiveness of the models used (Pollick, 2009).

Tinwell (2011) suggests that the Uncanny Valley problem may never be solved because the increasing proficiency of creating realistic digital characters is matched by users corresponding aptitude in evaluating digital characters. In other words, as users are exposed to more sophisticated CGI characters they become more astute observers of such characters. As such they will always be able to spot flaws, thus hitting what Tinwell (2011) refers to as the "Uncanny Wall."

Even with studies like these taken into consideration, there isn't a great deal of hard data to support the idea that realism and higher levels of detail positively correlate with user acceptance. Most studies using digital characters have included the use of still images instead of animation, and many of these explicitly state that their findings may not apply to animated characters (Green et al.,2008; Rozin and Fallon, 1987; Schneider et al., 2007). Brenton et al. (2005) suggests that digital characters will appear less "uncanny" as users become accustomed to them.

A 2001 study indicated that the human face is so detailed and people are so adept at interpreting faces that a high level of facial detail could unintentionally produce misinterpretations of affective messages (Donath, 2001). In fact, there is evidence to suggest that stylized/ simplified characters do a better job of expression and communication. One research study indicated that simple drawings of body parts depicting gestures were generally easier to recognize than more complex representations (Godenschweger et al 1997).

In this case, the idea follows the observation that stylized characters have the ability to exaggerate features and movements that allow emotional cues to be more easily perceived. Examples of potential advantages are plentiful. For example, characters can be designed with bigger eyes to better perceive gaze and subtle changes such as twitching and pupil dilation. In fact all facial features can be enlarged to facilitate the perception of affective display. Other characters can have longer limbs to better facilitate expressive movement.

In simplest terms, stylized characters have the potential to be more expressive and better at affective display than "realistic" characters. For example, consider the characters of Benjamin Button from the 2008 film *The Curious Case of Benjamin Button* and Carl Frederickson from Pixar's 2009 animated film *UP (See Figure 13).*

Figure 13: Benjamin Button on the left and Carl Frederickson on the right. Which face do we respond to more?

Like most Pixar creations, Carl is a stylized version of a real-world entity (in this case a grumpy old man.) While very detailed, his features wouldn't be considered "realistic." While materials such as his clothing, skin and hair are very detailed, the general form of the character is simple and cartoon-like. Benjamin Button on the other hand, much more closely resembles a real person.

Is one character better suited for expression or emotion than the other? This is difficult to determine as so many other factors play into the emotional effectiveness of characters. Script, storyline, direction, acting, and the technical abilities of CG artists are just some of the factors at work when determining whether such characters work. Even then, these considerations don't even take into account the immense variability and preferences of the audience members themselves. (While these examples obviously consist of pre-rendered characters in film and animation, there is no reason to believe that many of these points won't translate to interactive characters.)

There is literary evidence to suggest that too much graphic detail can be a hindrance to communication. As previously mentioned, all communication models have the potential to fail if there is too much information being communicated. With characters such as Carl, it's possible that there could be a greater chance for emotional connection due to the simplicity of their design. The prevalence of stylized animated features produced by studios as opposed to the minimal number of realistic *Final Fantasy*-style features would seem to provide at least anecdotal evidence in this regard.

Characters such as *Benjamin Button* (while much more realistic than Carl) could possibly be considered too detailed even considering the fact that character was intended to be composited with live-action footage. Instead of focusing on the performance of the character, audience members could potentially be distracted by the detail of the graphics. Of course, there are those who would argue that if the CG technology can be implemented so that users/ audience members are unable to distinguish between actor and a CG constructs, the question of level of detail becomes irrelevant.

Level of detail and the theory of the Uncanny Valley are fascinating topics that are certainly relevant to the design of digital characters. But for the moment, it may have to sufficient to say that the discussion is still ongoing. As such, the question to how "real" digital characters should be will continue to be answered on a case-by-case basis, taking into account the purpose, function, and environment for which a character will be used.

Works Cited:

"Affect." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. 9 August 2011. Web. 01 August. 2011. <http://en.wikipedia.org/wiki/Affect_(psychology)>

"Affect Display." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. 16 January 2011. Web. 01 August. 2011. <http://en.wikipedia.org/wiki/Affect_display>

"Affect Theory." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. 24 February 2011. Web. 01 August. 2011. <http://en.wikipedia.org/wiki/Affect_theory>

"Affective Science." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. 11 January 2011. Web. 01 August. 2011. <http://en.wikipedia.org/wiki/Affect_(psychology)>

"Affective Computing." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. 9 August 2011. Web. 01 August. 2011. <http://en.wikipedia.org/wiki/Affective_computing>

Argyle, M. (1988) *Bodily Communication (second edition)*, New York, Methuen & Co. Inc

Argyle, M. (1994) *The Psychology of Interpersonal Behaviour (Fifth edition)*, London, Penguin Books

"Artificial Intelligence." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. 23 August 2011. Web. 01 August. 2011. <http://en.wikipedia.org/wiki/Artificial_Intelligence>

Bartlett, M.S. (1998) *Face Image Analysis by Unsupervised Learning and Redundancy Reduction*, PhD Thesis, University of California, San Diego

Benford, S.D., Bowers, J., Fahlén, L.E., Greenhalgh, C.M., Snowdon, D. (1995) User Embodiment in Collaborative Virtual Environments, in *Proceedings of 1995 ACM Conference on Human Factors in Computing Systems (CHI'95)*, Denver/Colorado, ACM Press

Bailenson, J., Yee, N. (2005). Digital Chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments, *Psychological Science*

Breazeal, C. (2003) Towards sociable robots, in T. Fong (ed) *Robotics and Autonomous Systems*, 42(3-4), 167-175

Carnegie, D. (1936) *How to Win Friends and Influence People*, New York, New York, Pocket Books

Cooper, B., Brna, P., Martins, A. (2000) Effective Affective in Intelligent Systems – Building on Evidence of Empathy in Teaching and Learning, in Ana Paiva (Ed.) *Affective Interactions: Towards a New Generation of Computer Interfaces*, Springer Verlag, pp. 21-34

Damásio, A.R. (1994) *Descarte's Error: Emotion, Reason and the Human Brain*, Avon Books, New York

Darwin, Charles (1872). The Expression of Emotions in Man and Animals.

Descartes, R. (1641) *Meditations and Other Metaphysical Writings*, Reprint, Penguin

Dellaert, F., Polizin, t., and Waibel, A., (1996) "Recognizing Emotion in Speech", In Proc. Of ICSLP 1996, Philadelphia, PA, pp.1970-1973

Dittrich, W.H., Troscianko, T., Lea, S.E.G., Morgan, D. (1996) Perception of emotion from dynamic point-light displays presented in dance, in *Perception*, vol 25, pp 727-738

Donath, J. (2001) Mediated Faces, in M. Beynon, C.L. Nehaniv, K. Dautenhahn (eds.), *Cognitive Technology: Instruments of Mind*, Proceedings of the 4th International Conference on Cognitive Technology, Warwick, UK, August 2001

Ekman, P., Friesen, W.V. (1975) *Unmasking the Face,* New Jersey, Prentice-Hall Inc.

Ekman, P., Friesen, W.V. (1978) *Facial Action Coding* System, Consulting Psychologists Press Inc.

Ekman, P., Rosenzweig, L. (1998) *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System,* Oxford University Press, ISBN 0195104463

"Emotion." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. 25 August 2011. Web. 01 August. 2011. <http://en.wikipedia.org/wiki/Emotion>

"Embodied Agent." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. 25 August 2011. Web. 01 August. 2011. <http://en.wikipedia.org/wiki/Embodied_agent>

Fogg B.J. (2003) Persuasive Technology: Using Computers to Change What We Think and Do, San Francisco, CA, Morgan Kaufmann

Green, R.D., MacDorman, K.F., Ho, C-C. and Vasudevan, S.K. (2008) 'Sensitivity to the proportions of faces that vary in human likeness', *Computers in Human Behavior*, Vol. 24,No. 5, pp.2456–2474.

Garau, M., Slater, M., Pertaub, D.P., Razzaque, S. (2005) The Responses of People to Virtual Humans in an Immersive Virtual Environment, in *Presence: Teleoperators and Virtual Environments*, 14(1), MIT Press, 104–116

Knapp, M.L. (1978) *Nonverbal Communication in Human Interaction (2*nd Edition) Holt, Rinehart and Winston Inc., New York

Lee, C.M.; Narayanan, S.; Pieraccini, R., (2001) Recognition of Negative Emotion in the Human Speech Signals, Workshop on Auto. Speech Recognition and Understanding

Lisetti C.L., Schiano, D.J. (2000) Facial Expression Recognition: Where Human-Computer Interaction, Artificial Intelligence and Cognitive Science Intersect, in *Pragmatics and Cognition (Special Issue on Facial Information Processing)*, vol 8(1), pp. 185-235

MacDorman, K.F., Ishiguro, H.(2006) The uncanny advantage of using androids in social and cognitive science research. Interaction Studies 7(3)

Morris, D., Collett, P., Marsh, P., O'Shaughnessy, M. (1979) *Gestures, their Origin and Distribution*, London, Jonathan Cape Ltd.

Myers, David G. (2004) "Theories of Emotion." Psychology: Seventh Edition, New York, NY: Worth Publishers, p. 500.

Nathanson, Donald L. (1992), Shame and Pride: Affect, Sex, and the Birth of the Self, New York: W.W. Norton

Norman D.A. (2004) Emotional Design: Why We Love (or Hate) Everyday Things, New York, NY: Basic Books

Persson P., Laaksoiahti J. , and L¨onnqvist P. (2000) Anthropomorphism - a multi-layered phenomenon. In K. Dautenhahn, editor, Proceedings of AAAI Fall symposium, Socially Intelligent Agents: The Human in the Loop, pages 131– 135. AAAI Press

Piaget J. (1933) A Handbook of Child Psychology, chapter Children's philosophies, pages 505–516. Worcester, MA, USA: Clark University Press, 2nd edition

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

Picard, R. (1999) Towards Interfaces that Recognise and Respond to a User's Emotional Expression, presented at *Affective Computing: The Role of Emotion in Human Computer Interaction*,British HCI Group Meeting, April 1999, University College London, UK

Pioggia, G., Hanson, D. Dinelli, S., Di Francesco, F., Francesconi, R., De Rossi, D. (2002) The Importance of Nonverbal Expression to the Emergence of Emotive Artificial Intelligence, in Yoseph Bar-Cohen (ed.) *Proceedings of SPIE,* Vol. 4695

Poggi, I., Pelachaud, C. (2000) Emotional Meaning and Expression in Animated Faces, in Ana

Pollick, F. (2009) In Search of the Uncanny Valley. Analog communication: Evolution, brain mechanisms, dynamics, simulation. Cambridge, MA: MIT Press: The Vienna Series in Theoretical Biology

Paiva (Ed.) *Affective Interactions: Towards a New Generation of Computer Interfaces*, Lecture Notes in Artificial Intelligence 1814, ISBN 3-540-41520-3, Springer Verlag, 182-195

Reeves B. and Nass C. (1996) The media equation: How people treat computers, television, and new media like real people and places. Cambridge, MA: Cambridge University Press

Rozin, P. and Fallon, A.E. (1987) ‘A perspective on disgust’, *Psychological Review*, Vol. 94, No. 3, pp.23–41.

Scassellati B. (2000) Discriminating Animate from Inanimate Visual Stimuli. In Proceedings of the 17th International Joint Conference on Artificial Intelligence, pages 1405–1410. San Francisco, CA, USA: Morgan Kaufmann

Schmitz, M. (2011) Concepts for Life-Like Interactive Objects, in TEI '11**:** Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction, New York, New York, ACM Press

Schneider, E., Wang, Y. and Yang, S. (2007) ‘Exploring the Uncanny Valley with Japanese video game characters’, *Proceedings of the DiGRA 2007 Conference*, Tokyo, Japan, pp.546–549.

Strongman, K.T. (1996) *The Psychology of Emotion. Theories of Emotion in Perspective (Fourth Edition)*, New York, Wiley & Sons

Tinwell, A., Grimshaw, M.and Williams, A. (2011) ‘The Uncanny Wall’, *Int. J. Arts and Technology*, Vol. 4, No. 3, pp.326–341.

Terzopoulos, D., Waters, K. (1993) Analysis and synthesis of facial image sequences using physical and anatomical models, in *Pattern Analysis and Machine Intelligence*, vol 15(6), pp. 569-579

"The Emotion Machine." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. 15 November 2010. Web. 01 June. 2011. <http://en.wikipedia.org/wiki/Affect_(psychology)>

Wasinger R. and Wahlster W. (2005) True Visions: Tales on the Realization of Ambient Intelligence, chapter The Anthropomorphized Product Shelf: Symmetric Multimodal Human-Environment Interaction, pages 291–306. Berlin, Heidelberg, Germany: Springer

White P. (1995) The understanding of causation and the production of action: From infancy to adulthood. London, UK: Psychology Press

William, James (1884). "What is Emotion". Mind 9: 188–205.

Yacoob, Y., Davis, L. (1994) Computing spatio-temporal representations of human faces, in *Proceedings of the Computer Vision and Pattern Recognition Conference*, p70-75, IEEE Computer Society

Zebrowitz, L.A. (1997) *Reading Faces: Window to the Soul?*, Boulder, Colorado, Westview Press