

Affective Tutoring Systems: Enhancing e-learning with the emotional awareness of a human tutor

Nik Thompson, Murdoch University, Australia*

Tanya Jane McGill, Murdoch University, Australia

ABSTRACT

This paper introduces the field of affective computing, and the benefits that can be realized by enhancing e-learning applications with the ability to detect and respond to emotions experienced by the learner. Affective computing has potential benefits for all areas of computing where the computer replaces or mediates face to face communication. The particular relevance of affective computing to e-learning, due to the complex interplay between emotions and the learning process, is considered along with the need for new theories of learning that incorporate affect. Some of the potential means for inferring users' affective state are also reviewed. These can be broadly categorized into methods that involve the user's input, and methods that acquire the information independent of any user input. This latter category is of particular interest as these approaches have the potential for more natural and unobtrusive implementation, and it includes techniques such as analysis of vocal patterns, facial expressions or physiological state. The paper concludes with a review of prominent affective tutoring systems and promotes future directions for e-learning that capitalize on the strengths of affective computing.

Keywords: e-learning, affective computing, intelligent tutoring system, cognitive-affective theory, affective tutoring system, learner emotions

INTRODUCTION

Affective computing is defined as 'computing that relates to, arises from, or deliberately influences emotions' (Picard, 1997, p. 3). Affective computer interfaces improve human-computer interaction by enabling the communication of the user's emotional state. While research in human-computer interaction in the past had been dominated by cognitive theories, the importance of users' affective response is gaining attention (e.g. Beale & Peter, 2008; Gratch & Marsella, in press; Scherer, Banziger, & Roesch, 2010). An important future step in interface design is to incorporate the findings of the body of affective computing research into interaction environments that enhance both cognitive performance and personal comfort by providing the needed emotional context (Maxwell, 2002). This is even more relevant given the shift from the desktop paradigm toward ubiquitous computing. As the computing environment is steadily becoming more tightly integrated with the day to day physical world, developments in this area are applicable to a vast array of situations such as embedded applications, information appliances, vehicles and so forth.

There is evidence that emotion has an impact on the speed at which information is processed (Öhman, 2001) and whether it is attended to (Anderson, 2001; Vuilleumier, 2001). Emotion also has a relation to motivation in that evaluations or feelings regarding the current situation will

largely determine the action that is taken in response. Therefore, emotions are often precursors of motivations (e.g. Oatley, 1992). Memory is also impacted by emotional state, and again there are many mechanisms by which this can occur. The Processing Efficiency theory (Eysenck & Calvo, 1992) suggests that emotions can utilize cognitive resources that would otherwise be used for processing new information; for example in the case of anxiety, intrusive thoughts may compete with the cognitive task and result in a decrease in performance. Thus, an area which can benefit greatly from affective computing is education. The fact that interaction with computers is a fundamental part of study in most disciplines, coupled with the cognitive and emotional journey that all learners experience makes e-learning an ideal candidate for affective computing developments.

Intelligent tutoring systems attempt to emulate a human tutor by providing customized feedback or instruction to students. Whilst intelligent tutoring systems remain an active area of research, they have failed to achieve widespread uptake. A reason for this is the technical difficulty inherent in building cognitive models of learners and facilitating human-like communications (Reeves, 1998). The difference in learning performance between ideal one-to-one tutoring conditions and other methods is known as the 2 Sigma problem (Bloom, 1984). Research on expert human tutors indicates that 'expert human tutors devote at least as much time and attention to the achievement of affective and emotional goals in tutoring, as they do to the achievement of the sorts of cognitive and informational goals that dominant and characterize traditional computer based tutors' (Lepper & Chabay, 1988, p. 242). Given the apparent link between cognition and affect, it may be argued that for an intelligent tutoring system to emulate a human tutor successfully there should be some consideration of affective processes during learning. The inability of current intelligent tutoring systems to cater for the role of emotion in learning may to some extent explain the 2 Sigma problem in the context of computer based learning. It is hoped that the incorporation of affective components into e-learning development may therefore lead directly to improved pedagogical outcomes. Providing this vital form of affective feedback into intelligent tutoring and other applications should greatly improve their success.

This paper supports a move toward affective computer based learning by reviewing the relevant underlying theories and the methods for sensing affect and then linking these with prominent affective tutoring systems currently under development. It is hoped that this paper will foster an awareness of the significance of this aspect of human learning, and stimulate development of richer e-learning experiences to maximize students' learning outcomes.

COGNITIVE BASIS FOR LEARNING

The past few decades have seen the rise of the personal computer to fill many varied roles as organizer, communicator, entertainer and of course, educator. Research in the area of learning has predominantly taken a cognitive view in which the mental processes are considered as they are involved in learning. Cognitive theory is a learning theory of psychology that attempts to explain human behavior by understanding the thought processes. Cognitive theory is based on the assumption that human beings are logical and will make rational choices.

The field of cognitive psychology provides explanations for many of the underlying mental processes that occur during learning. Prominent in this field is the three stage information

processing model (Atkinson & Shiffrin, 1968) shown in Figure 1. This multi-store model of memory proposes that incoming information from the environment is briefly captured in sensory memory, and that information that is interesting is more likely to go on from sensory memory to short term memory. If a particular piece of information needs to be retained, the learner then makes a conscious decision to work with it and to continue to process it. Information that the learner has deemed important is eventually encoded to the long term memory for storage and later retrieval.

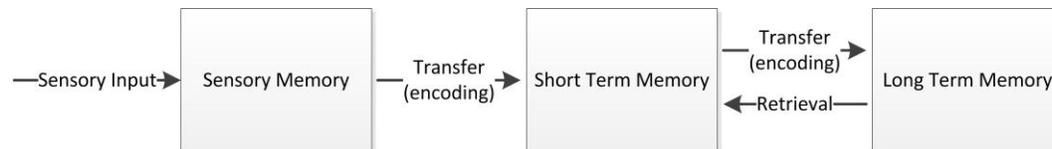


Figure 1. Three Stage Information Processing Model (Atkinson & Shiffrin, 1968)

More recently constructivism has gained ground; constructivists believe that learners' reality is built upon their existing experiences and perceptions. What someone knows is grounded in perception of the physical and social experiences which are comprehended by the mind (Jonassen, 1991). However, in spite of the way in which learning theories may have evolved over time, they have shared the perspective that the human mind is viewed as an information processing tool, not unlike basic computer architecture.

Perkins highlighted the compatibility between traditional cognitive theories and constructivism, stating '...information processing models have spawned the computer model of the mind as an information processor. Constructivism has added that this information processor must be seen as not just shuffling data, but wielding it flexibly during learning -- making hypotheses, testing tentative interpretations, and so on.' (Perkins, 1992, p. 51).

Cognitive theories however do not explain the role that emotions play, in spite of the substantial evidence that emotions influence cognitive processes (Pekrun, 2008). Norman (1981) cited the topic of emotion as one of the major challenges to cognitive theory. Some authors consider the information-processing metaphor as the source of this challenge; for example, Ortony, Collins and Clore (1990, p. 5) stated 'This approach to cognition has been as noticeable in its failure to make progress on problems of affect as it has been for its success in making progress on problems of cognition'.

People cannot be viewed purely as task-solving, goal driven agents, they also have other emotive reasons for their choices and behavior that drive the decision making process (Mandler, 1975). Lisetti (1999) claims that a large number of cognitive tasks are influenced by affective state, including organization of memory, attention, perception and learning. The same conclusion was reached by Picard (1997, p. x) who states that 'emotions play an essential role in rational decision making, perception, learning and a variety of other cognitive functions'.

COGNITIVE-AFFECTIVE THEORY

Another important area of research considers the underlying affective or emotional states and how these interact with cognitive processes. The way in which affective states interact with memory, decision making and social behavior creates a challenge for cognitive theory (Andrade & May, 2004). Emotions may disrupt, slow down, organize or initiate cognitive processes, and different emotions can influence these mechanisms in different ways (Pekrun, 2002). There has been a strong bias toward the cognitive and rational within the field of computer science, as a result of the prevailing view that the sciences are the domain of rules and logic with little room for anything else (Picard, 1997). In this view, emotion would be considered more of a distraction than a benefit. This bias has been reflected in the development of e-learning software, as it would generally be developed by programmers rather than learning theorists or educators. Consequently, many of the benefits of research into human affect and emotion are not yet fully realized in e-learning software.

In the field of e-learning, a popular theory describing how learners process and learn from computer based multimedia is Mayer's (2001) Cognitive Theory of Multimedia Learning. This theory draws from the multi-store model of memory described above, and others, to form a unified theory of the various aspects of cognitive processing of multimedia content and provides guidelines for instructional developers to improve learning outcomes. Central to the theory are the concepts that the human cognitive processes include limited working capacity, dual channels for various types of material (sound/images) and that the information is actively processed and assimilated by the learner (Mayer, 2001). Moreno (2006) extended this model to include the role of affect in learning and named it the Cognitive-Affective Theory of Learning with Media (see Figure 2). Where it differs from the original model is in the inclusion of affective and motivational factors. This addition acknowledges the role of affect as a mediator for rational cognitive processes such as learning.

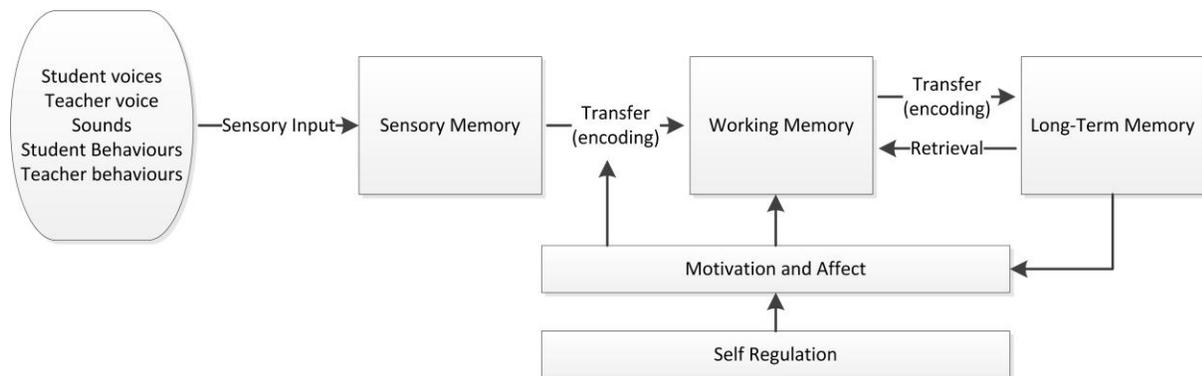


Figure 2. Cognitive-Affective Theory of Learning with Media (Moreno, 2006)

According to this theory, the level of interest that the learner has in the material will correlate to learning benefits by influencing students to invest more effort in the task. Furthermore, some instructional methods may be more supportive than others therefore producing improved learning outcomes by improving the student's feelings about their ability to complete the task (Moreno, 2006). The author discusses the effect of emotions such as anxiety or confidence, but this theory could potentially also apply to a wider range of more subtle emotional expressions.

Cognitive psychologists are not the only ones recognizing the link between emotion and mental processes; emotion theorists have long recognized that emotion itself may have a cognitive component. Schacter and Singer (1962) are known for their 2-factor theory in which they argue that there is a cognitive determinant to emotion. Before this work, emotion was believed to reflect biologically determined responses, and this perspective evolved to the view that emotion was a consequence of cognitive process and that various external factors determine the emotion that would be felt (Andrade & May, 2004). What this implies is that cognition and emotion are deeply intertwined, and that future developments in affective applications must acknowledge the two way interaction between these two basic areas of human functioning.

THE ROLE OF AFFECT IN LEARNING

Stein and Levine (1991) have identified a link between a person’s goals and emotions, and proposed a goal-directed, problem solving model. As with other theories of emotion that indicate that people like to maximize positive affective states, their model assumes that people attempt to assimilate information into their existing knowledge – when this information is new it results in arousal of the autonomic nervous system – this, in conjunction with a cognitive appraisal results in an emotional reaction. Therefore this model predicts that learning always occurs during an emotional episode.

Kort, Reilly and Picard (2001) have developed a model that links emotions and stages of learning in a four quadrant spiral (see Figure 3). The learning process is broken up by two axes, vertical and horizontal to signify learning and affect. The learning axis contains labels to indicate a range from constructive learning at one end, to un-learning at the other. The affect axis ranges from negative to positive. When a learner is working through a task with ease, they will be in quadrant I, experiencing constructive learning and positive affect. As the material becomes harder or if they struggle, they would move through quadrants II, III and finally IV. At this point they may be uncertain how to progress, but as they acquire new insights and ideas they will ultimately progress back to quadrant I so that the spiral may continue as they acquire more knowledge.

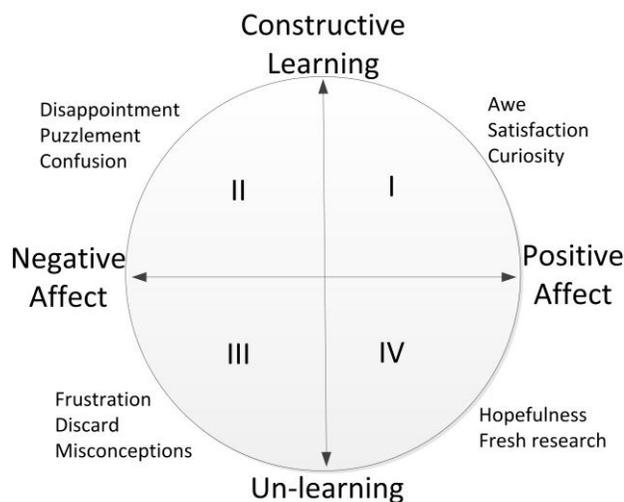


Figure 3. Model relating phases of learning to emotions (Kort, et al., 2001)

Goleman (1995) reported that expert teachers are able to recognize emotional states of students, and respond appropriately to positively impact learning. Whilst the way in which this is accomplished is not well documented, and may indeed differ between teachers, the foundation is still the same: to recognize negative affect or states that are detrimental to learning and to guide the learner into a more positive and constructive state. Csíkszentmihályi (1990) described an ideal learning state, which he called the zone of flow. In this state, time and fatigue disappear as the learner is absorbed and immersed in the task they are undertaking. When in a state of flow, people are absorbed in the activity and feel in control of the task and environment (Hsu & Lu, 2004). These characteristics of flow, are identical to what players experience when immersed and fully engaged in games (Chen, 2007), indeed games which create a flow experience are likely to be adopted, whilst others are discarded (Sherry, 2004). Thus educational games may also benefit from this effect, as the engagement and enjoyment of the learner is a catalyst to mediate their future learning and interest (Fu, Su, & Yu, 2009).

Intelligent tutoring systems attempt to emulate the personalized instruction that a human teacher may provide by building an internal model of the students' knowledge, abilities and progress. An e-learning system with these characteristics can have many advantages; for example being always available and potentially being able to provide more individual attention than in a traditional class based lesson. Intelligent tutoring systems incorporating an emotional or affective model are known as affective tutoring systems. An affective tutoring system is thus any tutoring system that can adapt to perceived emotion. This may be to respond to any negative emotions being experienced by the learner, or to interact in a manner that is more natural and engaging for the learner. These systems have also been shown to be effective and result in increased learning (as compared experimentally to a non-affect sensing implementation), however are still not as effective as a one-to-one human tutor. Further work is required.

For theories linking learning and affective states to be implemented into the development of affective tutoring systems an important consideration is the means by which the affective state can be inferred by the computer. The next section discusses the options that are available, and this is followed by a review of affective tutoring systems.

INFERRING THE LEARNER'S AFFECTIVE STATE

Given a suitable model to map affective states to desired behaviors or outcomes, the (technological) challenge is how to detect or infer the emotional state of the learner in the first place. There are several approaches to this, each with their own strengths and shortcomings.

One of the key issues surrounding the inference of affective state is the relationship between the underlying emotion and the observable expression or behavior which accompanies it. Schachter (1962) argued that the differentiation of emotion is not physical, but cognitive, and the data does support the fact that various observable signals may be common to a multitude of differing emotional states. Some signals are better than others for differentiating affective states, and one point which is agreed upon is that no single signal is a sufficient indicator of emotional response (Picard, 1997).

Affective states are internal and involve cognitive processes and are therefore not directly accessible to anyone other than the one experiencing them. Therefore it is only the observable

manifestation of the affective state that may be used for the process of inference. This is where the subtle, non-verbal indicators of underlying affect become especially useful. A further question is whether emotions may be categorized into discrete states, or whether they are dimensional constructs, which vary along a continuum with several components. According to discrete emotion theories, certain emotions like happiness, fear, sadness or interest are considered to be discrete, unique states that are experienced as the result of distinct causes (e.g. Izard, 1977); Many discrete emotion theories share the idea that a specific set of emotions is more basic or primary than the other emotions. These emotions are related to action tendencies and will thus have a physiological referent. In dimensional models of emotions, it is assumed that emotions can be represented in terms of a number of component dimensions (e.g. Russell, 1980). This viewpoint has the benefit of removing the need to categorize emotional experience within pre-defined boundaries, and may thus allow for a more fine-grained level of description.

Self Report

A multitude of self-report measures have been developed and used in research on mood and emotion; many of these share similar features but also differ in the way that the items are formatted, the instructions used and variations in the descriptive terminology applied. Many of the most prominent affective measures involve presenting lists of adjectives to the subjects, and obtaining a rating on a 4 or 5 point scale as to how appropriate or strong these particular emotions are (examples include: MAACL (Zuckerman & Lubin, 1965) , POMS (McNair, 1971) or PANAS (Watson, Clark, & Tellegan, 1988)). Depending on the test in use, the questions may refer to the current day, previous week or general overall emotional state.

More recently developed, the Current Mood Questionnaire (CMQ) is a complex instrument that uses multiple response formats for several dimensions of affect (Feldman-Barrett & Russel, 1998; Yik, Russell, & Feldman-Barrett, 1999). Mood is assessed through several means: 1. simple adjectives rated on 5 point Likert scale; 2. more complex mood statements rated using an agree/disagree format; and 3. trait like descriptions rated on a 4 point scale.

Although the CMQ is generally considered to be internally consistent and reliable, and results are satisfactory for the pleasantness/unpleasantness dimension, the results are less than satisfactory for the measures of arousal or activation dimension. These problems are not unique to the CMQ and it has also proven difficult to create good measures of this dimension in other measuring instruments (Watson & Vaidya, 2003). Overall, since the CMQ is a rather time intensive method, it is not often used as a practical affect measuring instrument.

The use of self-report also introduces some specific challenges. In particular, since the subjects are being relied upon for their input, the success of the measurement depends on them being firstly aware of their own internal affective experiences and secondly to be able to accurately express these within the constraints of the assessment tool. The quality of self-report will be directly related to the ability of the subjects to accurately identify feelings, and for them to be asked the right questions, at the right time and in the best manner (Levenson, 1988). Due to the subjective nature of these judgments it can be argued that there is a considerable risk of errors, even unintentional, when using this method. Self-report measures are indeed subject to both random and systematic measurement errors (Coan & Allen, 2007).

Observable Traits

Emotions are said to produce ‘pervasive, although generally short-lived, changes in the organism as a whole (Scherer, 1995, p. 235). Thus, there are several aspects of emotional expression that are observable. The use of observations to infer the emotional state of an individual stems largely from the work of Ekman and colleagues who theorized relationships between particular facial configurations and the underlying emotions present. The Ekman, Friesen and Tomkins Facial Affect Scoring Technique (FAST) (1971) specified what they believed to be the distinctive components of six categories of affect expressions. This was based on previous research and was highly theoretical in nature. FAST, however could not be used to determine whether facial actions other than those specified are relevant to emotion. This theory was developed into the more widely known theory of ‘basic emotions’, in which Ekman theorized that there are a set of basic emotions (Ekman & Friesen, 1978). This theory was developed further to derive lists of facial expressions that would be used as markers for these emotions.

Ekman and Friesen’s Facial Action Coding System (FACS) was designed to measure all facial activity and not just actions related to emotion (Ekman & Friesen, 1978). However FACS is slow to learn and use and requires slow motion viewing of facial actions. It is therefore unsuitable for real time coding. A further issue with all measures of emotions which use observations is that of independent validation – a common approach in research is to ask subjects to report their feelings (retrospectively) and see whether the facial expressions differ from those expected, this technique brings with it the issues that are associated with the use of self-report as an assessment tool.

Although less frequently studied, there are other observable aspects of emotional expression. These include expressions such as posture or vocalization. Empirical support for the ability for listeners to successfully recognize emotional state from vocal cues has been provided in many studies spanning the last 50 years (e.g. Lieberman, 1961; Scherer, 1986; Williams, 1972). On average the reported accuracy is around 60%, which is substantially better than the (12%) result that would be obtained purely by guessing.

Psychophysiology

Researchers have become increasingly aware that a critical component of emotion is physiological activity. According to some theories, if there is no physiological reaction there is no emotion (e.g. Schachter & Singer, 1962). Often a multi-modal approach is taken, with the view that emotion involves a complex pattern of responses, in which physiology plays a role. This view is by no means a recent development; William James (1890) speculated that patterns of physiological response could be used to recognize emotion. It is theorized that every psychological event or affective state has some physiological referent (Cacioppo & Tassinary, 1990), therefore the issue is not so much of whether or not a physiological signal is present, but rather which aspects of emotion may be inferred from this signal.

There are vast arrays of physiological expressions which may be suitable for inferring affective state; these include easily measurable expressions such as muscle movement or breathing rate, to more subtle measures such as neural activation of muscles, brain activity, skin conductance and cardiovascular measures. There is empirical data linking patterns of physiological response to specific affective states, however results are mixed, and in some cases inconclusive (Cacioppo &

Tassinary, 1990). Therefore, the use of physiological measures brings with it a rich and varied resource of information about the individual, but possibly an equally substantial amount of data processing considerations regarding how to interpret the data. However, there are arguably many advantages to this approach. Physiological signals are unconscious and do not carry any of the subjectivity of self-report measures, furthermore they bring about the potential for real time measurement with no need to interrupt or otherwise distract the user. Finally, as technology advances, physiological sensors may be suitable for incorporating into existing physical interfaces to ensure a more natural interface which the user need not be constantly aware of.

AFFECTIVE TUTORING SYSTEMS

This section discusses the prominent affective tutoring applications that have been developed. The input mechanisms are discussed for each system as well as the domain, and possible future directions and improvements are discussed where appropriate.

AutoTutor is an intelligent tutoring system that interacts with learners using natural language and helps them to construct explanations in simulation environments (Graesser, McDaniel, & Jackson, 2007). The current version of AutoTutor detects the learner's affective state using physiological and facial expression analysis and conversational cues. The AutoTutor focuses on a model of learner's emotions that includes emotions such as boredom, engagement, confusion or delight. The responses given by the tutoring system are designed to regulate the occurrence of any negative emotions in the learner. Initial results indicate that the affective tutor improved learning (as compared to a non-affective implementation of AutoTutor), particularly for low domain knowledge learners (D'Mello, Lehman, & Graesser, 2011).

Other projects have also examined the prediction of emotions using conversational cues as opposed to physiological data. A successful example involving use of dialogue features is ITSPOKE (Litman & Silliman, 2004). ITSPOKE is a spoken dialogue system that uses the Why2-Atlas physics tutoring system as its back-end (VanLehn et al., 2002). The student begins by typing in a natural language answer to a physics problem, after which the ITSPOKE system engages the student in a spoken dialogue to elicit more information and clear up misconceptions.

In another project involving ITSPOKE, Litman and Forbes-Riley (2004) used dialogue features to predict human emotion in computer-human tutoring dialogues, and to provide the ability for the software to detect uncertainty on the part of the learner and respond to address this. Although no significant differences were observed in metrics of student performance, the automated emotion prediction did outperform the baseline in all cases, however was not as successful as emotion prediction by a human. They did establish the utility of using acoustic and lexical features to infer emotion, and this may be beneficial for applications which utilize this means of interaction.

Conati (2002) developed a probabilistic model to monitor a user's emotions and engagement during their interaction with educational games. The model incorporates aspects of user interface input and physiological markers to estimate their emotional state. The dependencies between emotional states and possible causes is based on a cognitive model of emotions (Ortony, et al., 1990). The model relies on a dynamic decision network to utilize indirect indicators of the users' emotional state. The goal being that the model may be used by pedagogic agents to guide the

timing and type of interactions that will occur with the user. To evaluate this model, the Prime Climb educational game developed at the University of British Columbia was used as a test bed. The game helps students to learn number factorization with a two player climbing game in which players must solve factorization problems to progress. The original game has a pedagogical agent which provides hints when prompted. The affective version of this game utilizes the model of learner's affect to guide the actions of the agent and also to select the appropriate affective expression to display. When tested with year 6, 7 and 8 students the authors found a significant difference in test scores between the affective and non-affective groups for the younger students, but did not observe significant results with the older year 7 and 8 students. They attributed this partly to a ceiling effect found whereby the older students had already mastered the topic, but further investigation is needed to establish the causes. These results are promising, given that the inference of learners affect in this model is probabilistic and based on the student's progress in the game – using a more direct measure of learner's affect would very likely yield a better classification rate and potentially better outcomes.

Woolf, Burelson and Arroyo (2007) have developed several methods to evaluate students' emotion using facial expression, skin conductance, posture and mouse pressure, using Bayesian networks, not unlike the models proposed by Conati. (2002). A number of experiments have been carried out to recognize and respond to emotions in a learning context. In one study, an on screen agent interacted with the learner when frustration was detected. The agent responded to frustration with empathetic or task-support dialogue. Results demonstrated that students became more motivated after receiving the feedback. In a similar study by some of the same authors, machine learning was used to estimate the student's engagement using measures for student proficiency, motivation, evidence of motivation and student's response to a problem. Their software used the measure of the student's engagement to predict the probability of a correct student response with up to 75% accuracy and showed that disengagement negatively correlates with performance gain (Johns & Woolf, 2006). In a further study Beal, Arroyo, Woolf, Murray and Walles (2004) modeled student affective characteristics using a mathematics tutor to guide the actions of the software in terms of interaction and hints given.

Easy with Eve is an affect sensitive mathematics tutor developed by the Next Generation Tutoring Systems project (Alexander, Sarrafzadeh, & Hill, 2006; Sarrafzadeh, Alexander, Dadgostar, Fan, & Bigdeli, 2008) at Massey University in New Zealand. Affect recognition is performed by video analysis to capture facial expression and gesture information from the user. The facial expression analysis builds upon the work of Ekman and Friesen's facial action coding system where various 'basic' emotions are described in terms of their facial movements. Facial features are extracted from the video input and a fuzzy facial expression classifier separates these into seven affective states. The inferences made about the user's affective state from this data are then utilized in a case based reasoning approach to dictate responses and behaviors of the animated on screen agent 'Eve'. The case based reasoning is reported to be slow due to the large amounts of data being processed; however the authors identified that this is an issue which could be addressed should the necessity arise.

Edu-Affe-Mikey is an affective tutoring system that features an animated agent tutoring in the domain of medicine. Affect inference is done by processing input from the keyboard and microphone. Human experts were consulted to develop a list of events which signify changes in

learners' emotional state. The occurrence of these events is detected, and a simple weighted average method is used to select the most likely emotional state which would result from such a combination of events. This information is then used to select one of the pre-programmed responses presented by the on screen animated agent (Alepis, Virvou, & Kabassi, 2008).

Predinger, Dohi, Wang, Mayer and Ishizuka (2004) have developed an Empathic Companion: an animated interface agent that detects and responds to the user's affective state. The software uses physiological signals of skin conductance and muscle movement to infer the emotional state in terms of its component dimensions. The agent is intended to address the user's emotional state by showing concern in the form of empathic behavior. As one of the aims is to make the interaction as natural as possible, this affect recognition process is done in real-time while the user is interacting with the computer. The software application is presented in the context of a job-application interview scenario, where the affective agent responds to emotions elicited by the interview process. This physiological data is not constantly processed; rather it is made available when interface events request it, for example at the end of each interview question. A Bayesian network is employed to decide the most likely emotional state based on the input data set and to select from a number of pre-defined animation sequences to be presented by the Microsoft Agent based on screen character. The study carried out using this Empathic Companion had some limitations. Due to technical limitations of the ProComp+ physiological data acquisition hardware it was not possible to record and process the physiological data simultaneously. A workaround was put in place with a second set of (non-identical) hardware. This technical limitation also meant that muscle movement sensing was not possible in the study, and this was substituted for with a more basic measure of heart rate. These issues with the hardware and implementation of the experimental setup could to some extent explain the lack of statistically significant results from the study. However, the authors also make a strong point that the nature of the interview task may not induce the kinds of emotions that can be measured by this method. The authors suggest that an Empathic Companion would be more suitable for use in computer based education.

CONCLUSION

This paper has presented an overview of the motivation for applying the benefits of affective computing to e-learning. It has introduced the field of affective computing, and the benefits that can be realized by enhancing e-learning applications with the ability to detect and respond to emotions experienced by the learner. In order to understand the potential value of affective computing for e-learning theories of learning and the role of affect in learning have been reviewed. Some of the potential means for inferring the affective state of learners were also considered. These can be broadly categorized into methods that involve the user's input, and methods that acquire the information independent of any user input. This latter category is of particular interest as these approaches have the potential for more natural and unobtrusive integration, and it includes techniques such as analysis of vocal patterns, facial expressions or physiological state. The paper also included a review of prominent affective tutoring systems that have been developed and discussed the approaches that have been taken in each of them.

The review of affective tutoring systems has brought several things to light. Firstly, it has highlighted that the approaches taken are all diverse, down to the detail of which kind of affective data is acquired, how it is processed, interpreted and even how the affective tutoring

system responds to this data. In fact, no two systems reviewed are alike. Secondly, it confirms that these systems are scarce. This is in spite of the clear support for the role of affect in learning that is provided by psychological theories, and evident in the results of the evaluations of affective tutoring systems that have been developed and tested to date. We believe that these two observations are linked and that the current requirement for ad-hoc development in affective computing is hampering progress. Allanson and Fairclough (2004) noted that research in the area was disparate and uneven, and it seems that little progress has been made since then.

The application of affective computing to learning, is a cross disciplinary area, drawing from diverse fields such as computer science, psychology and education. Thus, a successful development either requires a developer to possess expertise in several distinct areas, or to have the support of a large research group. This could contribute to the observed scarcity of affective tutoring systems in the literature. However, this requirement is not necessarily a weakness, but may rather be turned to the advantage of developers under the correct conditions. What is required is to abstract the functional components of an affective tutoring system into a generalizable and re-usable model which will allow developers to build upon their successes iteratively and incrementally. Such a framework or 'blueprint' for affective tutoring systems, will also facilitate modularization of solutions and allow separate groups to work on different functional components within their own area of expertise, thus eliminating the above mentioned issue associated with such cross disciplinary work. Further research is required to develop this model.

In the short term, interface designers and educators may still learn from the successes of affective tutoring systems and draw from the principles that were applied to their development. Educators should aspire to incorporate some level of affective enhancement into any educational applications. Even if the software is unable to 'read' the emotional or cognitive state of the learner, the evidence still stands that learning benefits can be obtained by maximizing positive affect. Cognitive theories such as Mayer's cognitive theory of multimedia learning (Mayer, 2001) have received widespread interest from educators and multimedia designers, and the application of cognitive principles to any multimedia lesson has been shown to improve learning. This benefit is observed even in software that does not possess an internal model of the learner's cognitive state (e.g. Thompson & McGill, 2008). In fact, these cognitive principles may be treated like 'best practices' and successfully applied by the developers of any educational materials. This is the era of affect in computing, and the next logical step is to develop affective theories of multimedia learning, to provide similar guidelines for how to present material in such a way as to maximize positive affect. This will enable all instructional developers to draw from the growing body of affective computing knowledge, and translate this into improved tutoring interfaces to the benefit of the learners.

This is an exciting time for e-learning – the worldwide e-learning sector generated \$32.1 billion in 2010, and has been growing at 9.2% per year over the last 5 years (Adkins, 2011). This growth should not be perceived as pressure to move the same content from physical to electronic delivery, but as an opportunity to dramatically re-design educational materials in line with these new insights into learning. Innovations that bring improved educational outcomes, whilst ensuring the scholastic, motivational and affective goals of the learner are balanced in a supportive and natural learning environment, should be embraced.

REFERENCES

- Adkins, S. S. (2011). *The worldwide market for self-paced elearning products and services: 2010-2015 forecast and analysis*: Ambient Insight Research LLC.
- Alepis, E., Virvou, M., & Kabassi, K. (2008). Requirements analysis and design of an affective bi-modal intelligent tutoring system: The case of keyboard and microphone. In M. Virvou & L. C. Jain (Eds.), *Intelligent Interactive Systems in Knowledge-Based Environments*.
- Alexander, S., Sarrafzadeh, A., & Hill, S. (2006). Easy with Eve: A functional affective tutoring system. In G. Rebolledo-Mendez & E. Martinez-Miron (Eds.), *Proceedings of Workshop on Motivational and Affective Issues in ITS. 8th International Conference on ITS 2006* (pp. 38-45).
- Allanson, J., & Fairclough, S. H. (2004). A research agenda for physiological computing. *Interacting with Computers*, 16, 857-878.
- Anderson, A. K. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, 411(6835), 305-309.
- Andrade, J., & May, J. (2004). *Cognitive Psychology*. Sheffield: BIOS Scientific Publishers.
- Atkinson, R. C., & Shiffrin, R. M. (1968). *Human memory: A proposed system and its control processes*. Oxford: Academic Press.
- Beal, C. R., Arroyo, I., Woolf, B. P., Murray, T., & Walles, R. (2004). Intelligent tutoring systems web-based intelligent multimedia tutoring for high stakes achievement tests. In J. C. Lester, R. M. Vicari & F. Paraguaçu (Eds.), *Proceedings of 7th international conference on intelligent tutoring systems : ITS 2004, Macei, Alagoas, Brazil, August 30-September 3, 2004* (pp. 468-477): Springer.
- Beale, R., & Peter, C. (2008). The role of affect and emotion in HCI. In P. Christian & B. Russell (Eds.), *Affect and Emotion in Human-Computer Interaction* (pp. 1-11): Springer-Verlag.
- Bloom, B. S. (1984). The 2 sigma Problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4-16.
- Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45(1), 16-28.
- Chen, J. (2007). Flow in games (and everything else). *Communications of the ACM*, 50(4), 31-34.
- Coan, J. A., & Allen, J. J. B. (Eds.). (2007). *Handbook of Emotion Elicitation and Assessment*: Oxford University Press.
- Conati, C. (2002). Probabilistic assessment of user's emotions in educational games. *Applied Artificial Intelligence*, 16(7-8), 555-575.
- Csikszentmihályi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.
- D'Mello, S., Lehman, B., & Graesser, A. (2011). A motivationally supportive affect-sensitive AutoTutor. In R. A. Calvo & S. K. D'Mello (Eds.), *New perspectives on affect and learning technologies* (Vol. 3, pp. 113-126): Springer New York.
- Ekman, P., & Friesen, W. (1978). Facial action coding system: A technique for the measurement of facial movement. *Consulting Psychologists Press*.
- Ekman, P., Friesen, W., & Tomkins, S. S. (1971). Facial affect scoring technique, (FAST) a first validity study. *Semiotica*, 3(1), 37-38.

- Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion*, 6(6), 409-434.
- Feldman-Barrett, L., & Russel, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, 74(4), 967-984.
- Fu, F., Su, R., & Yu, S. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers and Education*, 52(1), 101-112.
- Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.
- Graesser, A. C., McDaniel, B., & Jackson, G. T. (2007). Autotutor holds conversations with learners that are responsive to their cognitive and emotional states. *Educational Technology*, 47, 19-22.
- Gratch, J., & Marsella, S. (Eds.). (in press). *Social emotions in nature and artifact: Emotions in human and human-computer interaction*. New York: Oxford University Press.
- Hsu, C.-L., & Lu, H.-P. (2004). Why do people play on-line games? An extended TAM with social influences and flow experience. *Information and Management*, 41(7), 853-868.
- Izard, C. (Ed.). (1977). *Human Emotions*. Delaware: Plenum Press.
- James, W. (1890). *Principles of psychology*. New York: Holt.
- Johns, J., & Woolf, B. (2006). A dynamic mixture model to detect student motivation and proficiency *Proceedings of Twenty-First National Conference on Artificial Intelligence* (pp. 163-168). Menlo Park, California: AAAI Press.
- Jonassen, D. H. (1991). Objectivism versus constructivism: do we need a new philosophical paradigm? *Educational Technology Research and Development*, 39(3), 5-14.
- Kort, B., Reilly, R., & Picard, R. W. (2001, 6-8 Aug). *An affective model of interplay between emotions and learning: Reengineering educational pedagogy—Building a learning companion*. IEEE International Conference on Advanced Learning Technologies, Madison, USA.
- Lepper, M. R., & Chabay, R. W. (1988). Socialising the intelligent tutor: Bringing empathy to computer tutors. In H. Mandl & A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems* (pp. 242-257).
- Levenson, R. W. (1988). Emotion and the autonomic nervous system: A prospectus for research on autonomic specificity *Social psychophysiology and emotion: Theory and clinical applications*. (pp. 17-42): Oxford, England: John Wiley & Sons.
- Lieberman, P. (1961). Some aspects of fundamental frequency and emotional content in speech. *The Journal of the Acoustical Society of America*, 33(11), 1677.
- Lisetti, C. L. (1999). *A user model of emotion-cognition*. Proceedings of the First Workshop on Attitude, Personality, and Emotions in User-Adapted Interaction of the the International Conference on User-Modeling (UM'99), Banff, Canada.
- Litman, D. J., & Forbes-Riley, K. (2004). *Predicting student emotions in computer-human tutoring dialogues*. Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics, Barcelona, Spain.
- Litman, D. J., & Silliman, S. (2004). *ITSPOKE: An intelligent tutoring spoken dialogue system*. Proceedings of the Human Language Technology Conference 4th Meeting of the North American Chapter of the Association for Computational Linguistics, Boston, Massachusetts.
- Mandler, G. (1975). *Mind and emotion*. New York: NY:Wiley.

- Maxwell, K. (2002). The maturation of HCI: Moving beyond usability toward holistic interaction. In J. M. Carroll (Ed.), *Human-Computer Interaction in the New Millennium*. New York: ACM Press.
- Mayer, R. (2001). *Multimedia learning*: Cambridge University Press.
- McNair, D., Lorr, M., Dropplemen, L. (1971). *Edits manual: Profile of mood states*. San Diego: Educational and Industrial Testing Services.
- Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning*(22), 149-158.
- Moreno, R. (2006). When worked examples don't work. Is cognitive load theory at an impasse? *Learning and Instruction*, 16(170-181).
- Norman, D. A. (1981). Twelve issues for cognitive science. In D. A. Norman (Ed.), *Perspectives on cognitive science* (pp. 265—295,). Hillsdale, NJ: Erlbaum.
- Oatley, K. (1992). *Best laid schemes: The psychology of emotions*. Cambridge: Cambridge University Press.
- Öhman, A. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology*, 130(3), 466-478.
- Ortony, A., Clore, G. L., & Collins, A. (1990). *The Cognitive Structure of Emotions*: Cambridge University Press.
- Pekrun, R. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37(2), 91-105.
- Pekrun, R. (2008). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology*, 41(4), 359-376.
- Perkins, D. N. (1992). Technology meets constructivism: Do they make a marriage? In T. M. Duffy & D. H. Jonassen (Eds.), *Constructivism and the technology of instruction* (pp. 45-55). New Jersey: Lawrence Erlbaum and Associates.
- Picard, R. W. (1997). *Affective Computing*. Massachusetts: MIT Press.
- Prendinger, H., Dohi, H., Wang, H., Mayer, S., & Ishizuka, M. (2004). Empathic embodied interfaces: Addressing users' affective state. In E. André, L. Dybkjær, W. Minker & P. Heisterkamp (Eds.), *Affective Dialogue Systems* (Vol. 3068/2004, pp. 53-64): Springer Berlin / Heidelberg.
- Reeves, T. (1998). *The impact of media and technology in schools, A research report prepared for the Bertelsmann foundation*.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178.
- Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C., & Bigdeli, A. (2008). "How do you know that I don't understand?" A look at the future of intelligent tutoring systems. *Computers in Human Behaviour*, 24(4), 1342-1363.
- Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69, 379-399.
- Scherer, K. R. (1986). Voice, stress and emotion. In R. Trumbull & M. H. Appley (Eds.), *Dynamics of stress* (pp. 159-181). New York: Plenum Publishing.
- Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of voice*, 9(3), 235.
- Scherer, K. R., Banziger, T., & Roesch, E. (Eds.). (2010). *A blueprint for affective computing*. Oxford: Oxford University Press.
- Sherry, J. L. (2004). Flow and media enjoyment. *Communication Theory*, 14(4), 328-347.

- Stein, N. L., & Levine, L. J. (1991). Making sense out of emotion. In W. Kessen, A. Ortony & F. Kraik (Eds.), *Memories, thoughts, and emotions: Essays in honor of George Mandler* (pp. 295-322). Hillsdale, NJ: Erlbaum.
- Thompson, N., & McGill, T. J. (2008). Multimedia and cognition: Examining the effect of applying cognitive principles to the design of instructional materials. *Journal of Educational Computing Research*, 39(2), 143-159.
- VanLehn, K., Jordan, P. W., Ros, C. P., Bhembé, D., Bottner, M., Gaydos, A., et al. (2002). *The architecture of Why2-Atlas: A coach for qualitative physics essay writing*. Proceedings of the 6th International Conference on Intelligent Tutoring Systems.
- Vuilleumier, P. (2001). Beware and be aware: Capture of spatial attention by fear-related stimuli in neglect. *Neuroreport*, 12(6), 1119-1122.
- Watson, D., Clark, L. A., & Tellegan, A. (1988). Development and validation of brief measures of positive and negative affect : The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070.
- Watson, D., & Vaidya, J. (2003). Mood measurement: Current status and future directions. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology* (Vol. 2 : Research methods in psychology, pp. 351-375).
- Williams, C. E. (1972). Emotions and speech: Some acoustical correlates. *The Journal of the Acoustical Society of America*, 52(4b), 1238.
- Woolf, B., Burelson, W., & Arroyo, I. (2007). *Emotional intelligence for computer tutors*. Proceedings of 13th International Conference on Artificial Intelligence in Education.
- Yik, M. S. M., Russell, J., & Feldman-Barrett, L. (1999). Structure of self-reported current affect: Integration and beyond. *Journal of Personality and Social Psychology*, 77(3), 600-619.
- Zuckerman, M., & Lubin, B. (1965). *Manual for the multiple affective adjective check list*. San Diego: Educational and Industrial Testing Service.