Education Tech Research Dev
DOI 10.1007/s11423-016-9470-5

RESEARCH ARTICLE

Genetics with Jean: the design, development and evaluation of an affective tutoring system

Nik Thompson[1] · **Tanya Jane McGill**[2]

Abstract This paper details the design, development and evaluation of an affective tutoring system (ATS)—an e-learning system that detects and responds to the emotional states of the learner. Research into the development of ATS is an active and relatively new field, with many studies demonstrating promising results. However, there is often no practical way to apply these findings in real-world settings. The ATS described in this paper utilizes a generic affective application model to infer and appropriately respond to the learner's affective state. This approach brings several advantages, notably the potential direct support for re-use and retrospective addition of affect sensing functionality into existing e-learning software. Skin conductivity and heart rate variability measurements were used to infer affective activation and valence. The evaluation involved an experiment in which the effectiveness of the fully functional ATS was compared with that of a non-affective version, and was conducted with 40 adult participants. The evaluation of the effectiveness of this tutoring system showed that measurable improvements in perceived learning may be obtained with a modest level of software development.

Keywords Evaluation of CAL systems · Human–computer interface · Intelligent tutoring systems · Interactive learning environments

✉ Nik Thompson
nik.thompson@curtin.edu.au

Tanya Jane McGill
t.mcgill@murdoch.edu.au

1 School of Information Systems, Curtin University, Kent Street, Bentley, WA 6102, Australia

2 School of Engineering and Information Technology, Murdoch University, South Street, Murdoch, WA 6150, Australia

Published online: 05 August 2016

Introduction

Emotion and cognition are linked and there is evidence of emotion influencing aspects of cognitive performance and decision making (Cytowic 1989; Eysenck et al. 2007). Affect plays a role as a source of information, providing context and meaning to situations (Schwarz and Clore 1988) and directing attention and highlighting knowledge for further processing (Bower and Forgas 2001). Affect also functions as a motivator, influencing the tendency to approach or avoid a situation as well as how information is processed (Frijda 1986; Custers and Aarts 2005).

Emotion is thus strongly associated with whether information is attended to and how it is processed. Negative emotions are a widely agreed upon source of interference in the ability to process new information (Ellis and Ashbrook 1988). It has been suggested that this may be an adaptational characteristic to encourage slower and more systematic processing. Negative emotions are a source of information to indicate that a problem has occurred and that future processing should be carried out in a more systematic and rational (and thus slower) manner (Schwarz 1990). It is apparent that these effects of underlying emotional state may influence the processing and acquisition of new information during learning. As the learning experience is simultaneously influenced by emotional state, and also directly influences the emotional state of the learner—this relationship is both fundamental and complex.

These insights into the role that emotion plays gave rise to the proposition that a learning session may be improved if the teacher is sensitive and responsive to the emotional state of the learner. There exists a body of evidence to support this assertion in human–human tutoring (Goleman 1995) and this interest has now been extended to e-learning (D'Mello and Graesser 2012a). The central principle is generally similar: If a tutoring system were able to respond to the learner's emotional state, there would be measurably improved learning outcomes. A system that did this could incorporate insight into the learner's emotional state to guide the course of the lesson; for example to offer hints when confusion is detected, or to offer a break when boredom sets in. Research in the domain of affective tutoring systems (ATS) has yielded a number of developments. These systems often incorporate functionality including tutoring strategies, affect sensing and learner progress tracking into a single, highly integrated environment. Empirical evaluations of these ATS suggest that they can contribute to the learning experience (D'Mello et al. 2011; Conati 2002; Alexander et al. 2006; Wu et al. 2015). However there is often no practical way to directly apply these findings to a real learning setting in order to discover the optimum environment for ATS success. As such there is no widespread use of ATS outside the controlled research setting.

This paper reports on the design, development and evaluation of the effectiveness of an ATS called Genetics with Jean, in which the animated tutor, Jean, responds to the affective state of the learner. This ATS was built as a proof of concept for an approach intended to facilitate broader adoption of ATS, by enabling addition of affect support to existing e-learning applications. It uses a model for affective application development (Thompson and McGill 2015) and differs from existing systems due to the loose coupling between the tutoring and affect sensing functionality. This is advantageous as it provides direct support for re-use of software components outside the research environment and introduces the prospect of layering affect sensing functionality into existing, widely used e-learning applications.

Background

People prefer to be in certain states (e.g. pleasure) rather than others (e.g. pain): when unpleasant states are experienced, an attempt is made to regulate and to change these states toward more favourable ones by initiating the appropriate thinking or action. Stein and Levine (1991) propose that people continuously monitor their environment in this attempt to maximise positive states and when new information is detected this interrupts the otherwise routine pattern matching process and attention may shift to the new information. Thus emotional experience is believed to be associated with acquiring and processing new information, and learning is believed to always occur during an emotional episode.

The benefit of emotional awareness during human–human tutoring dialogues is already established; expert teachers respond to the emotional states of students, and guide their progress in a way that has a positive impact on learning (Goleman 1995). This appears to involve recognition of potentially detrimental negative states followed by guidance to a more positive state conducive to learning.

The role of emotion in multimedia learning systems has received some attention. In early work, Kort et al. (2001) developed a two dimensional model that linked emotions and stages of learning and proposed approaches to validating it. Craig et al. (2004) identified six main affective states during a learning session with a tutoring system and found a positive relationship both confusion and flow, and learning, and a negative one between boredom and learning. Moreno's (2006) Cognitive-Affective Theory of Learning with Media builds on Mayer's Cognitive Theory of Multimedia Learning (Mayer 2005) and takes into account both motivational and affective aspects of multimedia learning. Support for the Cognitive-Affective Theory of Learning with Media (Moreno 2006) has been provided in recent studies; for example, Plass and colleagues have shown that positive emotions generated during multimedia learning facilitate comprehension of material (Plass et al. 2014) and increase learners' mental effort, satisfaction and motivation (Um et al. 2012).

Research into affective user interfaces has demonstrated that timely provision of affect support via animated agents (e.g., Prendinger et al. 2004) or intelligent tutoring systems (e.g., Sarrafzadeh et al. 2008) can also alleviate negative feelings, and research that supports this is discussed below. These findings support the proposition that the incorporation of affective components into tutoring systems can result in enhanced learning outcomes.

Related work

There are a number of ATS described in the literature with each tending to use different input modalities, and tailored to work with specific instructional material. The internal architecture of these systems is often highly complex, and they are constructed very differently. However, in terms of high level functional components there do exist commonalities between different platforms: an ATS would naturally require some sort of affect sensing functionality (an *affective model*), there may also be a *domain model* to incorporate subject matter knowledge, and a *student model* to incorporate the knowledge about the learner to guide the tutor to respond appropriately. In spite of these high level commonalities, due to the lack of standard implementation architecture, it is not possible to directly compare approaches aside from concluding that each ATS is either successful or unsuccessful in its own way.

Some of the most prominent ATS include affect sensitive versions of AutoTutor (D'Mello et al. 2011; D'Mello and Graesser 2012a, 2012b), Prime Climb (Conati and Zhao 2004) and Easy with Eve (Alexander et al. 2006; Sarrafzadeh et al. 2008). The range of domains for which ATS have been designed is broad (Wu et al. 2015). Mathematics has been the most common domain with systems including Prime Climb (Hernández et al. 2008), Easy with Eve (Alexander et al. 2006; Sarrafzadeh et al. 2008) and Wayang (Woolf et al. 2009; Woolf et al. 2007), but ATS for database modelling (EER-Tutor; Zakharov et al. 2007), physics (ITSPOKE; Litman and Silliman 2004), medicine (Edu-Affe-Mikey; Alepis and Virvou 2011), computer literacy (D'Mello and Graesser 2012a), learning a foreign language (Lin et al. 2015), build environment (Kaklauskas et al. 2015) and job interview preparation (Empathetic Companion; Prendinger et al. 2004) have also been developed and trialled.

The types of input modalities used in ATS include facial expression analysis, conversional interaction with the interface, and physiological measures. Facial expression analysis has been used in many ATS including Affective AutoTutor (D'Mello et al. 2011), EER-Tutor (Zakharov et al. 2007), Easy with Eve (Alexander et al. 2006; Sarrafzadeh et al. 2008), Wayang (Woolf et al. 2009; Woolf et al. 2007), an implementation of an affective e-learning framework (EMASPEL) by Ammar et al. (2010), and a system by Lin et al. (2016). Analysis of dialogue features to predict emotion has been used in ITSPOKE (Litman and Silliman 2004) and in Edu-Affe-Mikey (Alepis et al. 2008).

Despite the shortcomings associated with the requirement for specialised equipment (Afzal and Robinson 2011), relative to other input modalities, physiological measures have several advantages: they may be obtained without interrupting the student from their task or disturbing their concentration and as technology advances, physiological sensors will be suitable for incorporating into existing physical interfaces to ensure a more natural interface that the user need not be constantly aware of. It is common to use several physiological measures in combination in the same ATS. For example, Prendinger et al. (2004) used skin conductance and muscle movement in their Empathic Companion, and Woolf and colleagues (Woolf et al. 2009; Woolf et al. 2007) used facial expression, skin conductance, posture and mouse pressure in Wayang. An even more extensive range was used in a prototype developed by Shen et al. (2009): heart rate, skin conductance, blood volume pressure, and EEG brainwaves.

While many prototype ATS have been described in the literature, only a limited number of them have had extensive evaluation of their impact on learning outcomes undertaken. The most extensive and encouraging results are those for Affective AutoTutor (D'Mello and Graesser 2012a), where use was associated with significant learning gains, measured via knowledge tests, for low domain knowledge college students. Affective versions of Prime Climb (Conati and Zhao 2004; Hernández et al. 2008) have also been shown to contribute to significant learning outcomes in primary school aged children.

Development goals and approach

The examples of ATS discussed above are indicative of the breadth and diversity of development in this domain. Each of these systems has several commonalities apart from their ability to detect and respond to emotions. Firstly, although implemented differently, the basic areas of functionality described in the previous section are always present—this includes a student model, an affective model, and a domain model. Secondly, each of these

ATS has been evaluated in some way. However a third, and less fortunate area of commonality also exists and that is the fact that there is no direct way to replicate and extend these systems, or for a non-expert to create new lesson material to be taught by these existing ATS. This is an issue that has been identified by developers of some existing ATS and has resulted in developments such as the AutoTutor lesson authoring tool (Susarla et al. 2003), however for the most part there does not yet exist a viable solution for all ATS platforms.

The Affective Stack Model (Thompson and McGill 2015) holds promise to provide a solution to this issue (see Fig. 1). It provides an architecture in which the main functionality of an ATS is described as a set of loosely coupled modules. The architecture is such that individual functional components of the affective computing system are not tied to a particular implementation and if used to guide ATS development, the system is not tied to a particular instructional context. This translates into a system that supports rapid development, incremental improvement and the ability to apply affective functionality to existing tutoring software.

The use of an architectural model such as the Affective Stack Model generally requires standardization of the functional components used in software. As described above, the common ATS include many of the same functional components: for example, an affective platform, which consists of both hardware and signal processing feature and a rule set that encodes the affective inputs and their associated responses for a given application context. These common components are incorporated into the Affective Stack Model. However, the ATS described in the literature consist of complex and highly coupled functional components. This complexity may sometimes be necessary to support the kind of complex and customized processing that takes place, however, a central question is whether a more

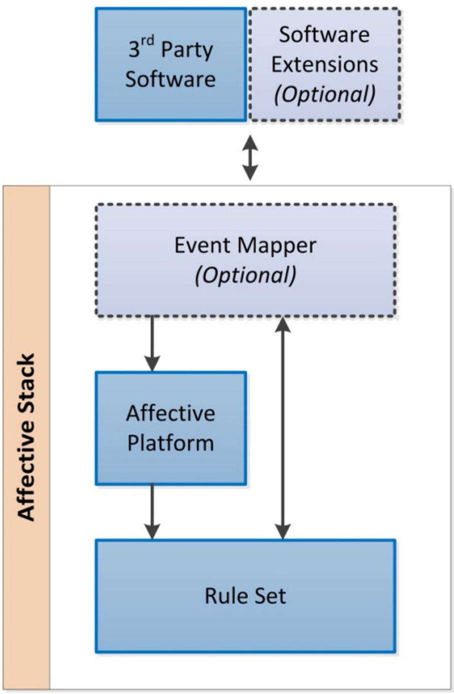

Fig. 1 Affective stack model (Thompson and McGill 2015)

streamlined and simplified application based on the Affective Stack Model is able to still yield positive outcomes. If this is possible, it will demonstrate the potential for existing tutoring software to be enhanced with affective capability.

This paper describes the design, development and evaluation of an ATS that was built on top of an existing set of web-based instructional materials, using the Affective Stack Model as a framework for development. The affective functionality was added to existing software and communication between components is via simple text files. It was built as a proof of concept for an approach intended to facilitate broader adoption of ATS, by enabling addition of affect support to existing e-learning applications. As this is a prototype system, there is scope for further enhancement in the future with the development of a more complex and rich feature set.

Genetics with Jean

Genetics with Jean is an ATS that teaches the subject of genetics. Information is presented in a number of modalities including text, graphics, diagrams and animation, which are all based on the Morgan Genetics Tutorial (Sofer and Gribbin 2010). This is a multimedia tutorial that covers the basic principles of genetics with a molecular slant, and was used with permission from the authors. The topic of genetics was considered relatively suitable for this proof of concept as it is a topic about which participants were unlikely to already possess a high level of knowledge (Molster et al. 2009; Richards 1996). It was also necessary that the tasks be sufficiently complex and engaging to elicit an affective response. The presence of in-text quizzes and questions in this set of materials contributed to this objective.

The learner's affective state is inferred by physiological signals using a previously developed affective platform, the details of which are available from Thompson et al. (2012). The underlying affective model uses a dimensional view of emotions (Prendinger et al. 2004), in which the affective state is considered in terms of the components of affective activation (the extent of activation or arousal experienced by the individual) and valence (how strongly positive or negative this experience is) (Prendinger et al. 2004). The physiological signals measured were selected to give insight into these two dimensions and to thus classify the likely affective state of the learner. The system is self-calibrating and signals are considered in the context of the learner's own recent range of physiological expression rather than comparison against a pre-selected 'baseline'. This insight into the affective state is then used to guide the actions of an on-screen animated pedagogical agent (APA), Jean. The purpose of the APA is to provide guidance and support to the user, and to emulate a human tutor. APAs have been shown to improve engagement and support learning (Kim et al. 2007; Moreno 2005) and have been widely used in ATS (e.g., D'Mello et al. 2011; Hernández et al. 2009). The agent, Jean, was based on the Microsoft Agent environment (Microsoft Corporation 2009) and has many animations, which may be scripted within software to achieve a believable and natural interaction. The Genetics with Jean software is a bespoke application, however the external components that provide the animated agent and the text to speech engine are freely available tools provided by Microsoft. A female character was chosen as research has demonstrated that a female character is more successful than a male character at reducing user frustration (Hone 2006). Speech balloons are accompanied by a spoken voice to make the agent more believable as a learning companion, as well as conforming to the principles of multimedia instruction that students learn better when explanations are provided auditorily rather than visually (Mayer 1998). See Fig. 2 for a sample screen from the ATS with the APA visible.

As noted previously, ATS are often highly complex and specialized software artefacts, utilizing precisely calibrated conditions and equipment. This ultimately limits the widespread applicability of the findings. Therefore, in this implementation, technology which may be conducive for later transference into consumer applications was chosen. Two physiological sensors were utilized to infer the dimensions of affective activation and valence. The sensors consist of small (3 mm) electrodes which simply require skin contact to operate. These were affixed to the skin with adhesive tape during the evaluation; however it is anticipated that in future refinements to the hardware they will be embedded into interface mechanisms such as a mouse as has previously been demonstrated (e.g. Kirsch 1997).

The emotional activation of the learner is inferred by measuring the conductivity of the skin. Skin conductivity tends to increase when a person is startled or experiences anxiety and is generally considered to be a good measure of overall level of activation (Cacioppo et al. 2007).

Emotional valence can be inferred using heart rate based measures. Whilst these are commonly measured using an electrocardiogram using electrodes on a chest strap this is quite an intrusive means of physiological measurement and detracts from its usefulness,

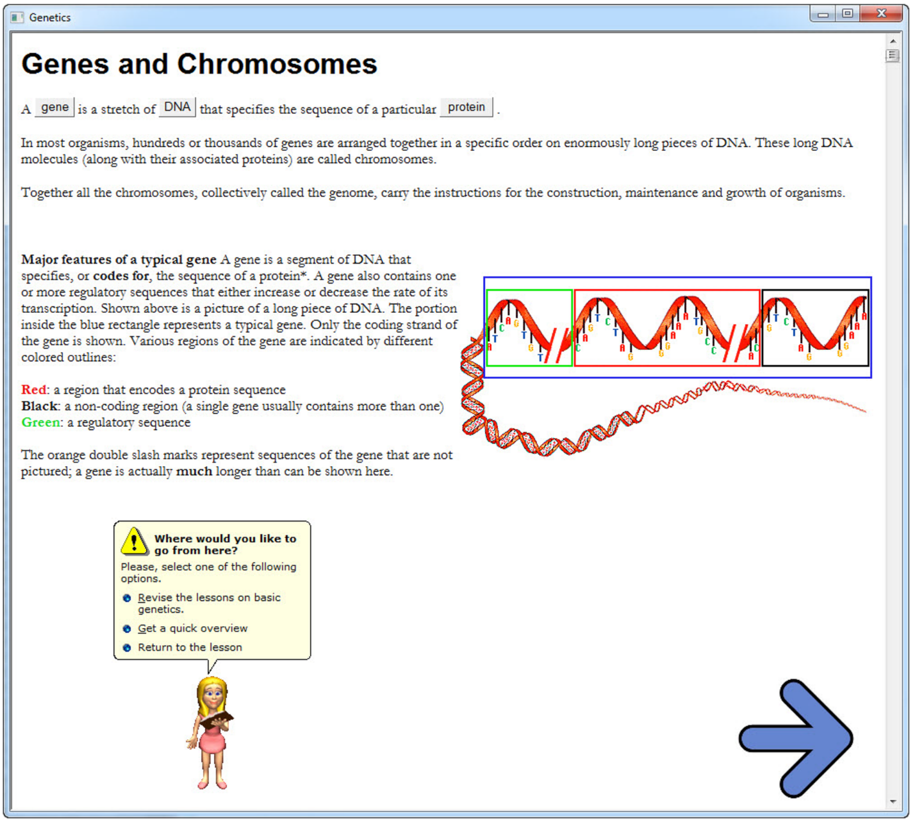

Fig. 2 Sample lesson screen

especially for an affective computing application (D'Mello 2008). Therefore in this implementation a more suitable approach was developed which simply requires skin contact using the same 3 mm sensors. Heart rate variability measures provide a rich source of information and make it possible to analyse specific frequency bands which correspond to certain underlying processes; for example, to discern between physiological changes due to physical exertion as opposed to affective state (Yannakakis et al. 2008). Therefore in this implementation heart rate variability was chosen as the measure of emotional valence.

Genetics with Jean operates in a fully automated real-time processing mode. This is a mandatory requirement in order for the system to function in its intended role as a real-time, interactive tutor. In keeping with the aim of promoting transferability, the reliance on commercial hardware platforms was kept minimal, with a generic National Instruments data acquisition device used to collect the physiological signals. This approach also made it possible to create a physiological data acquisition system that was superior in many ways to the commercially available products. This is particularly noteworthy in the implementation of heart rate variability measures. Commercially available platforms such as those available from ProComp and BioPac often carry certain limitations, the most apparent being the choice of physical sensors and form factors, as these devices are designed only to operate with the particular choice of hardware sensors that are provided by the manufacturer. A review of the current offerings from ProComp and BioPac also revealed that the analysis of heart rate variability measures was often only supported in off-line processing mode—a fact which alone renders them unsuitable for a real-time application such as an affective tutor. A further improvement over commercial platforms was in the responsivity of the system. Heart rate variability measures operate using blocks of data—the size of the block will influence how responsive the system is to detecting changes. Commercial platforms generally offer fixed pre-sets for data block size such as 3, 5 or 10 min (Thought Technology 2010) or allow the user to pre-select a block in offline mode (BioPac Systems 2004). In this implementation, signal pre-processing techniques were developed which make it possible to operate successfully using a much smaller block size of 100 s of data. The shorter block size for data acquisition directly correlates to the amount of time needed for the system to detect an emotional state change. This directly reduces the latency between the affective state being expressed and subsequently detected by the system, yielding a system that is more responsive in its operation.

In Genetics with Jean the animated agent serves as an affective tutor that aims to address, and alleviate, negative affective states so that they do not hinder the learning process. It does this by adopting an instructional strategy based on research on the role of emotional and motivational support in learning. A number of scripted behaviours were incorporated into Genetics with Jean, and these fall into the following three distinct modes:

- *Support Affect* The agent provides empathic feedback to the learner to address any potentially negative affective state; for example '*I know the material is quite hard*' or '*It's ok if you don't score full marks, you can always come back to this another time*'. This aspect of the support to learners is based on research showing that reducing negative emotions experienced by the learner via supportive messages can improve learning and satisfaction (Klein et al. 2002; Kim 2012; Shen et al. 2009).
- *Motivate* The agent provides supportive comments to keep the learner on track; this may simply be '*Well done*', or '*Keep up the good work*'. If the lesson is nearing the end, then the agent may also indicate that the task is almost complete if the student is losing interest. This motivational support was included consistent with research indicating the value of motivational messages in instructional software (van der Meij 2013; Kim 2012).

- *Offer Revision* The agent provides extra tutoring for topics that have triggered a negative affective state in the learner or when the learner's performance on quizzes is poor. This may take the form of separate revision pages with a summary of a whole topic, or simply links to enable them to easily look back to re-check content on a previous page.

The classification of affective state based on activation/valence dimensions was derived from that of Prendinger et al. (2004) and the following general rules were applied:

1. No action taken if affective state is neutral.
2. Offer Revision triggered if lesson performance is low.
3. Support Affect triggered if negative state detected (this is characterized by higher activation and negative valence).
4. Motivate response given if high activation detected with positive valence. This could indicate frustration or simply distraction. This response is generally motivational and does not address any specific state.

The on-screen agent behaviour is guided by two evaluations. The Instructional Module evaluates lesson performance and triggers the on-screen agent in situations where lesson performance (from quizzes) is low. This module is always enabled and operates for all users regardless of affective state. The next evaluation is carried out by the Affective Module, in which the affective activation and valence are interpreted to guide the actions of the on-screen agent in accordance with the general rules listed above. The Affective Module is optional and may be disabled if required. Therefore, in an experimental situation it is feasible to disable the Affective Module in a control group in order to be able to evaluate its efficacy. Figure 3 shows the affect processing rules that are applied within the Affective Module.

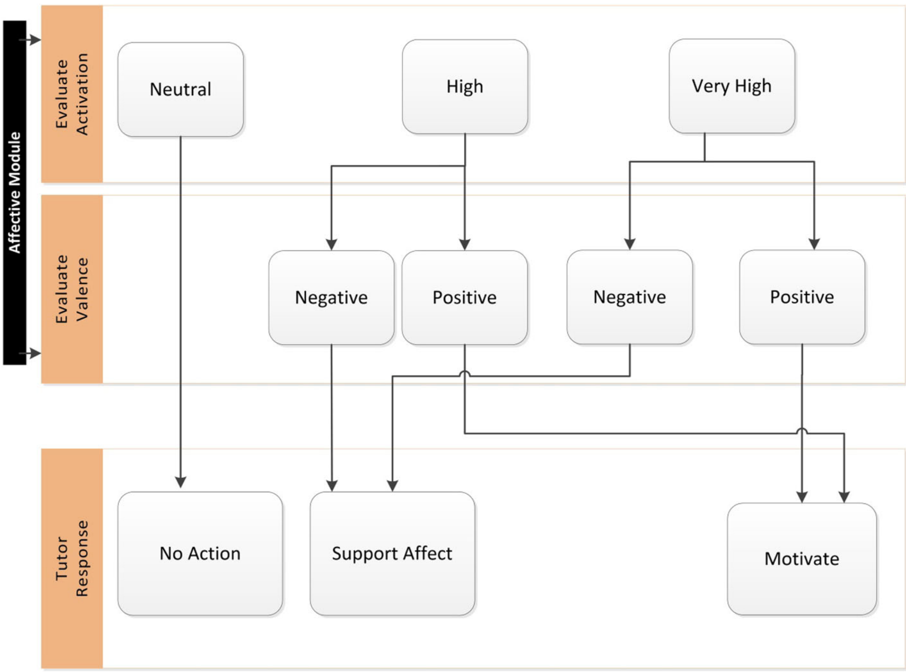

Fig. 3 Affective module states and actions

Before the formal evaluation was undertaken, piloting was carried out with five participants. This piloting targeted both the affective platform and Genetics with Jean. During this piloting, a manual analysis of the data logs from the software and affective components was undertaken in order to confirm that the software was behaving as expected. The raw data logs from the affective activation and valence sensors were first plotted on a time-series line graph. Next, the software event logs were overlaid on this graph. This data included timestamps for quizzes, lesson performance and any occasion where the on-screen agent was triggered. This visualisation enabled the researcher to pick any time point x and inspect the affective valence and activation at this point and observe what action was taken by the system at the time. This allowed for some fine-tuning of the hardware environment and sensors and made it possible to compare the expected action of the system with the actual behaviour and confirm correct operation before allowing the formal evaluation to proceed.

Evaluation of the genetics with Jean ATS

An empirical evaluation was carried out to determine the effectiveness of Genetics with Jean in improving outcomes associated with use of an ATS. The evaluation focused on the contribution of affective support to the effectiveness of the system.

Effectiveness, in an e-learning usability context, is broadly defined as the attainment of instructional objectives or that the learning outcomes are consistent with the learners' expectations (Pajares 1996). For the purposes of this study, effectiveness is considered in terms of three dimensions: Content Knowledge, Perceived Learning and Enjoyment. Content Knowledge is defined as the extent to which facts about the lesson content are retained by the learner. Perceived Learning is defined as self-reported perceptions of learning accomplishments (Jiang 2000). Enjoyment refers to the extent to which the activity of using the system is perceived to be enjoyable in its own right, apart from any performance considerations that may be anticipated (Malone 1981), and is thus considered to be a form of intrinsic motivation (Davis et al. 1992).

It has long been claimed that student learning is influenced by affective state, including organization of memory, attention, perception and learning (Lisetti 1999; Picard 1997), and the evidence to support this is increasing. Fostering and supporting positive emotions in multimedia learning has been shown to be associated with increased student learning (Mayer and Estrella 2014; Plass et al. 2014), and with increased mental effort, decreased perceived difficulty, increased motivation and increased satisfaction (Um et al. 2012). Some ATS implementation have also been demonstrated to contribute to learning gains (e.g., D'Mello and Graesser 2012a; Conati and Zhao 2004; Hernández et al. 2008). Therefore students whose learning is supported by this affective e-learning software could be expected to retain more knowledge of the content than students who do not receive this affective support. This leads to the first hypothesis that:

H1 Students who receive affective support from Genetics with Jean will have higher levels of Content Knowledge than students who do not receive affective support.

In practice, learning a new skill or knowledge is a complex process, characterized by many internal processes, which may not be immediately observable or measurable. Perceived learning is the set of beliefs and feelings one has regarding the learning that has occurred. It can include both cognitive and socio-emotional perceptions and studies have found that these emotional perceptions are informative (Clore et al. 2001). It has been argued that using just test scores to operationalize learning may not always provide the best measure as direct academic performance is not the only indication that learning has occurred (Caspi and Blau 2008). Therefore, in this study an additional measure of learning was also considered and it is proposed that students whose learning is supported by affective e-learning should have a higher level of Perceived Learning than students who do not receive this support. This is consistent with studies that have shown that fostering positive emotions during learning sessions contributes to increased mental effort, decreased perceived difficulty, increased motivation and increased satisfaction (Um et al. 2012) and leads to the second hypothesis that:

H2 Students who receive affective support from Genetics with Jean will have higher levels of Perceived Learning than students who do not receive affective support.

Increased enjoyment and engagement have implications for learning, but their roles are not well understood (Lin et al. 2012). Perceived enjoyment has been shown to influence learners intentions to use internet-based learning media (Lee et al. 2005) and 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 and Lu 2004). Thus the enjoyment of the learner is believed to be a catalyst to mediate their future learning and interest (Fu et al. 2009), and Lin et al. (2012) have proposed guidelines for increasing enjoyment during informal learning that include mood building. An ATS responds to the learner's affective state and attempts to provide feedback and support to address any negative states that may arise, therefore students who receive the feedback and support from an ATS could be expected to find the lesson more enjoyable than those who did not receive this feedback and support. This leads to the third hypothesis that:

H3 Students who receive affective support from Genetics with Jean will have higher levels of Enjoyment than students who do not receive affective support.

Evaluation methodology

Data was gathered during a laboratory experiment in which 40 adult learners interacted with the Genetics with Jean ATS. To achieve the objective of evaluating whether the fully functional Genetics with Jean ATS would be more effective than a non-affective version, participants were randomly assigned into two groups. The affective group interacted with the full ATS, whereas for the non-affective group, affective components were disabled.

Measurement instruments

As described above, the overall effectiveness of Genetics with Jean was considered in terms of three dimensions: Content Knowledge, Perceived Learning and Enjoyment. The study utilized a post-test only design for the effectiveness dimension of Content Knowledge. Content Knowledge was measured using a quiz conducted at the end of the data collection session. Due to the nature of the tasks and quizzes, any pre-test performance measures might affect the participation levels later shown by participants. Any difficulties experienced during the pre-test may lead them to attend to the subsequent materials in a different way. This situation was described by Campbell et al. (1963), who stated that "while the pre-test is a concept deeply embedded in the thinking of research workers in education and psychology, it is not actually essential to true experimental designs" (p.25). The quiz consisted of pre-existing questions from the Morgan Genetics Tutorial (Sofer and Gribbin 2010), which were based on the content covered in the lesson. It consisted of 12 questions worth one mark each; these were a mixture of fill in the blank and true/false questions. The total quiz score for each participant was calculated by adding the individual marks giving a maximum attainable quiz score of 12. The quiz questions are detailed in Table 1.

There are aspects of the learning experience that may not necessarily be reflected in test scores, but that can be captured in a broader measure. Perceived Learning was measured using five items, adapted from Alavi et al. (2002). These items were measured on a five point Likert scale, ranging from strongly disagree (1) to strongly agree (5). The items are shown in Table 2. Reliability testing was conducted to ensure that the items used to measure Perceived Learning demonstrated sufficient internal consistency. Cronbach's alpha for Perceived Learning was 0.7, and the scale was thus found to be reliable (Nunnally 1978). An overall score for Perceived Learning was calculated as the average of the five items.

Enjoyment was measured using four items on a five point semantic differential scale. The items were adapted from Ghani and Deshpande (1994) and are shown in Table 3. Cronbach's alpha for Enjoyment was found to be 0.67. Whilst this is slightly lower than the desired threshold this was deemed to be acceptable for an exploratory study (Nunnally 1978). An overall score for Perceived Enjoyment was calculated as the average of the four items.

Table 1 Content knowledge quiz questions

DNA consists of ___ of genes arranged in a specific order
The genes which carry the sequence of particular proteins is specified in a stretch of ___
The long pieces of DNA are called chromosomes. T/F
Human beings have tens of thousands of different kinds of proteins, and each one has a specific function. T/F
Each species has the same number of chromosomes. T/F
During meiosis the number of chromosomes gets reduced by a factor of ___
The choosing of chromosomes to be passed on during meiosis is a random process. T/F
Sex cells or gametes contain half of the number of chromosomes of other cells. T/F
A die is rolled, find the probability that an even number is obtained. ___ out of ___
Two coins are tossed, the probability that two heads are obtained. ___ out of ___
A card is drawn at random from a deck of 52 cards. Find the probability of getting a queen. ___ out of ___
A jar contains 3 red marbles, 7 green marbles and 10 white marbles. If a marble is drawn from the jar at random, what is the probability that this marble is white? ___ out of ___

Table 2 Perceived learning items

I became more interested in the subject of genetics
I gained a good understanding of the subject of genetics
I learned to identify central areas in the subject
I was stimulated to do additional study in the area of genetics
I found the current lesson to be a good learning experience

Table 3 Enjoyment items

I found the lesson		
Interesting	………………..	Not interesting
Fun	………………..	Not fun
Exciting	………………..	Not exciting
Enjoyable	………………..	Not enjoyable

Participants

The 40 participants for this evaluation were recruited from staff and students of two Australian universities. The primary selection criteria were that participants must be in good overall health and not be taking any medication or drugs (including caffeine) that could influence their physiological state.

Potential participants were invited to take part in the study via email distribution lists or referred by previous participants. All participants were sent a copy of an information letter so that they would fully understand the requirements before committing to taking part. Prior to commencement of the recruiting process, ethics approval was sought and obtained.

Data collection session

Separate data collection sessions were carried out for each participant. These were held in a quiet computer laboratory with only the participant and researcher present to minimize distractions. The physiological sensors were attached to two fingers of the non-dominant hand, to ensure that their presence did not interfere as the participants interacted with the software. As all participants exhibit a slightly different range of physiological responses, the Genetics with Jean ATS performs a self-calibration during operation. To ensure that this was completed, the sensors were attached at least 2 min before the start of the tasks to ensure that sufficient data had been recorded for the software to reliably calibrate.

Participants were randomly assigned to either the affective group or the non-affective group using a software randomizer—the affective group interacted with the fully functional ATS, the non-affective group used identical software, but with the affective components disabled. The physiological sensors were attached in the same way to participants in both groups to ensure that they were treated consistently. The use of the Affective Stack Model to guide the development of the Genetics with Jean ATS meant that system components were not necessarily dependent on one another. Therefore, the ATS could operate correctly even in the absence of any physiological input. This characteristic streamlined the way in which the experiment was conducted as it meant that all participants could be treated in the same way.

Genetics with Jean interacts with the user via, Jean, an on-screen animated agent intended to emulate a human tutor. In order to be able to ensure that any observed differences could be attributed to the affective components of the software and not simply the presence of an animated agent, the animated agent was present for both groups. In the version of the software used by the non-affective group the animated agent was unable to detect affective cues and thus relied only on non-affective sources of input. This was limited to reporting on the student's progress in the lesson, and presenting revision topics if the student requested help.

The sequence of steps for the study and details of the activities undertaken in the data collection sessions are shown in Table 4. The study procedure was first explained and participants were able to ask questions before providing consent. After this, the measuring equipment was set up and the sensors were attached. The measuring equipment to be used in the experiment was left running during the remainder of the introductory discussion and software demonstration to allow the participants to become familiar with the setup.

The software was demonstrated to ensure that participants were aware of the navigation controls and the general screen layout of the lesson. While this was mostly self-explanatory it helped participants to feel comfortable when using the interface.

Part 2 of the session required the user to interact with the ATS to complete several short lessons and quiz questions. In the affective group, the behaviour of the on-screen animated agent was determined by the affective responses displayed by the participant. The ATS was able to sense and respond to any negative affective states being experienced by the participants. It was not considered suitable to impose a fixed time constraint on the learning session, as this may have increased the cognitive load on the learner as well as potentially inducing a negative state due to the "test-like" conditions during the quiz section (Wine 1971; Eysenck and Calvo 1992) Therefore, approximately 25 min was allocated to this, but participants could progress at their own rate.

The final stage of the data collection session comprised of completion of a questionnaire, which was filled out electronically and covered the three aspects of effectiveness in the following order: Perceived Learning, Enjoyment and Content Knowledge. Participants were also able to include other feedback. The Perceived Learning and Enjoyment items involve the participant's self-report of their internal state or experiences. Therefore it was desirable to present these items as soon after the lesson as possible while the experiences were fresh in their memory. Furthermore this presentation order also precluded the possibility of any negative affect experienced when answering the Content Knowledge quiz from interfering with the measurement of Enjoyment or Perceived Learning.

Results

A total of adult 40 participants took part in the study (22 female and 18 male), with the majority (55 %) being aged between 20 and 29. They were randomly distributed into the two groups, and no significant differences in gender distribution ($\chi^2(1, N = 40) = 1.62$; $p = 0.204$) or age distribution ($\chi^2(4, N = 40) = 3.96$; $p = 0.412$) were found.

Does affective support lead to improvements in students' knowledge?

A difference in Content Knowledge between the affective group and the non-affective group would suggest that affective support provided by Genetics with Jean leads to improvements in students' knowledge of the lesson content. Table 5 provides descriptive statistics relating

Table 4 Activities undertaken during the data collection session

Part	Activities	Approx. duration
1	Introduce study and obtain participant's consent. Demonstrate software	10 min
2	Interact with Genetics with Jean	25 min
3	Questionnaire	10 min

Table 5 Content knowledge group statistics

Group	N	Min	Max	Mean	SD
Non-affective	20	6.0	12.0	10.25	1.74
Affective	20	8.0	12.0	10.60	1.23

to Content Knowledge. Content Knowledge scores were distributed normally and an independent samples *t* test was considered suitable to test for differences between the groups. The results of the independent samples t-test indicated that there was no significant difference in levels of Content Knowledge between the affective ($M = 10.60$, $SD = 1.23$) and non-affective group ($M = 10.25$, $SD = 1.74$); t (38) = -0.73, $p = 0.47$, d = 0.23. Therefore the hypothesis that affective support increases acquisition of knowledge about the subject area, genetics, was not supported.

Does affective support lead to improvements in students' Perceived Learning?

Descriptive statistics about Perceived Learning are presented in Table 6 below. The Perceived Learning data did not meet the assumption of normality required for an independent samples t-test, therefore a Mann–Whitney *U* test was used to determine whether affective support leads to improvements in Perceived Learning.

Participants in the affective group reported significantly greater levels of Perceived Learning than those in the non-affective group (Mdn 4.00 vs 3.40; $U = 94.0$, $Z = -2.90$, $p = 0.004$, r = 0.46) and the hypothesis that students who receive affective support from Genetics with Jean will have higher levels of Perceived Learning was thus supported.

Does affective support lead to improvements in students' enjoyment of learning?

Descriptive statistics for Enjoyment are shown in Table 7. The enjoyment data did not meet the assumption of normality, therefore a Mann–Whitney *U* test was conducted to

Table 6 Perceived learning group statistics

Group	N	Min	Max	Mdn	SD
Non-affective	20	2.60	4.40	3.40	0.40
Affective	20	2.80	5.00	4.00	0.50

Table 7 Enjoyment group statistics

Group	N	Min	Max	Mdn	SD
Non-affective	20	3.00	4.50	3.75	.47
Affective	20	3.00	5.00	4.00	.68

evaluate the hypothesis that students who receive affective support from Genetics with Jean ATS will have higher levels of Enjoyment than students who did not receive affective support. This hypothesis was not, however, supported with the results indicating that there was no significant difference in the levels of Enjoyment between the affective and non-affective groups (Mdn 4.00 vs 3.75; $U = 148.50$, $Z = -1.40$, $p = 0.16$, r $= 0.22$).

Discussion

This paper reports on the design, development and evaluation of an ATS called Genetics with Jean. The system teaches introductory genetics, interacting with the user via an APA. An internal decision network utilizes data from the affective platform to infer the dimensions of affective activation and valence in real-time while the learner interacts with the system. It has the capability to perform various types of interaction including providing affect support to address negative states, encouraging the learner and providing revision and review tips.

This ATS was developed to address limitations of previous ATS: firstly that there is often no practical way to implement these systems in real-world settings for long term use; and secondly that it has not been easy to add affect support to existing tutoring systems. Genetics with Jean differs from previous ATS in that it used the Affective Stack (Thompson and McGill 2015) as a framework for development. The approach taken provides a way forward from the current ad hoc nature of ATS development.

The use of a component based model, such as the Affective Stack Model (Thompson and McGill 2015), brings together all of the functional units of an affective computing environment in an architecture that is compatible with third party software applications, enables developers to incrementally improve separate components, and supports the re-use of functional components for new applications. The use of the Affective Stack Model in the development of Genetics with Jean resulted in loose coupling between the tutoring and affect sensing functionality, which is advantageous as it provides direct support for re-use of software components outside the research environment and introduces the prospect of layering affect sensing functionality into existing, widespread e-learning environments. The potential of this approach was demonstrated by the way in which an existing tutoring application, the Morgan Genetics Tutorial (Sofer and Gribbin 2010), was built upon, illustrating the potential for adding affect sensing functionality to existing e-learning applications.

The effectiveness of Genetics with Jean in terms of supporting student learning was evaluated in terms of Content Knowledge, Perceived Learning and Enjoyment. Although the average levels for all three dimensions of effectiveness were higher for those students receiving affective support from the ATS, these differences were only significant for Perceived Learning. It is possible that the sample size was not sufficient to detect small improvements in Content Knowledge and Enjoyment, therefore the potential of Genetics with Jean to influence these dimensions of effectiveness should be further explored in future studies.

The improvement in Perceived Learning is consistent with that found by Alexander (2007) who noted that students who interacted with his affective tutor had marginally higher levels of perceived learning than those who interacted with a non-affect sensing version.

The lack of significant improvement in Content Knowledge, as measured by a summary quiz, may result from the levels of existing domain knowledge of the participants. The levels of Content Knowledge were relatively high for both groups, suggesting that a ceiling effect may have obscured potential differences in performance. Using a pre-test to determine the levels of existing genetics knowledge of participants would have allowed further insight this potential explanation of the results. It is possible that this ceiling effect may not have occurred with more challenging material or a different assessment technique and thus any differences would be clearer. This is consistent with the findings of D'Mello et al. (2011) who note that the affective implementation of their AutoTutor software is particularly appropriate for low domain-knowledge learners. Aist et al. (2002) also made a similar observation that including affective enhancements in a tutoring system improved students persistence with the task, but not their memory of facts. This implies that the benefits may be more apparent in the longer term rather than after a short evaluation. This should be investigated in future research.

Levels of Enjoyment were also not found to significantly differ in response to affective support. The reason for this may lie in the nature of enjoyment of learning. Ghani and Deshpande (1994) noted that task challenge and sense of being in control are key factors that result in enjoyment. The role of these factors may also explain the results in this study. Both groups of learners interacted with an identical set of materials and hence should have experienced a similar level of sense of control and task challenge, possibly explaining the similar levels of Enjoyment. Another factor potentially indicating Enjoyment is the time spent on task and prior studies have shown that time on task is a predictor of e-learning success (Brown 2001). Although time spent was not explicitly measured during the study, it was noted that almost all of the participants chose to interact with the ATS for longer than the anticipated study duration and comments such as the following describe the kind of flow experience that is sought in human–computer interaction: '*I'm tired out after that lesson, but I didn't realize I'd been concentrating on it for 45 min!*'. It is possible that this immersion and lack of awareness of time contributed to Enjoyment for both groups. However, given that the median Enjoyment score was higher for the affective group, further research to clarify the role of affect support in enjoyment of ATS may be warranted. While, for pedagogical reasons a time limit was not imposed on the lesson, future research may include detailed user-experience metrics including time spent on tasks and reading time which may be included in the research to reveal if this variable plays a moderating role on other metrics.

Conclusion

Emotion is recognized as an important factor influencing learning. ATS attempt to improve learning by providing affect support to students as part of the learning experience. The study described in this paper involved the development and evaluation of an ATS called Genetics with Jean. This system was developed with the goal of being proof of concept for a development approach that should facilitate the enhancement of existing e-learning systems to provide affective support. This kind of approach is needed to help address the lack of widespread adoption of ATS outside research environments. Genetics with Jean was shown to improve students' perceptions of their learning, but not immediately tested knowledge of the content. The results support previously reported observations that affective tutors are perhaps most useful in situations with low-domain knowledge learners (D'Mello et al. 2011).

Increased pressures and time-constraints on individuals' work and personal life require that teaching and learning techniques become more flexible, effective and efficient in order to remain viable. The consistent growth in adoption of e-learning demonstrates the value placed on flexibility and ease of access. Therefore advanced technologies such as ATS, which combine the ease of access and flexibility of e-learning with the effectiveness of human tutoring, hold promise and are consequently attracting attention within the industry (Lowendahl 2012). However, unless the field progresses to a point where it is possible to rapidly develop and reuse affective technological solutions, widespread uptake may never occur.

Developments in affective computing applications are generally highly implementation specific. As affective applications are considerably specialized and complex, there has to date been little discussion regarding the concept of repurposing these applications. Therefore, although a great deal of time may be invested in the development of an affective interface for an e-learning application, the complexity of later refining this application or porting it to a new implementation domain may ultimately be prohibitive. The development approach of utilizing a standard model of core functional components has shown potential as a way forward for future work in the domain of ATS. This paper has detailed the design, development and evaluation of one such ATS that was developed using a standardized and conceptually straightforward architecture. The findings and areas for future investigation have laid the foundations for a promising direction for research and practice.

References

Afzal, S., & Robinson, P. (2011). Designing for automatic affect inference in learning environments. *Educational Technology & Society, 14*(4), 21–34.

Aist, G., Kort, B., Reilly, R., Mostow, J., & Picard, R. (2002). Experimentally augmenting an intelligent tutoring system with human-supplied capabilities: Adding human-provided emotional scaffolding to an automated reading tutor that listens. In 4th IEEE International Conference on Multimodal Interfaces (pp. 483-490). Pittsburgh, PA, USA: IEEE Computer Society.

Alavi, M., Marakas, G., & Yoo, Y. (2002). A comparative study of distributed learning environments on learning outcomes. *Information Systems Research, 13*(4), 404–415. doi:10.1287/isre.13.4.404.72.

Alepis, E., & Virvou, M. (2011). Automatic generation of emotions in tutoring agents for affective e-learning in medical education. *Expert Systems with Applications, 38*(8), 9840–9847. doi:10.1016/j.eswa.2011.02.021.

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 Knowedge-Based Environments*. Berlin: Springer-Verlag.

Alexander, S. (2007). *An affect-sensitive intelligent tutoring system with an animated pedagogical agent that adapts to human emotion*. Albany: Massey University.

Alexander, S., Sarrafzadeh, A., & Hill, S. (2006). Easy with Eve: A functional affective tutoring system. In G. Rebolledo-Mendez, & E. Martinez-Miron (Ed.), Proceedings of Workshop on Motivational and Affective Issues in ITS. 8th International Conference on ITS (pp. 38-45).

Ammar, M. B., Neji, M., Alimi, A. M., & Gouardères, G. (2010). The affective tutoring system. *Expert Systems with Applications, 37*(4), 3013–3023.

BioPac Systems (2004). Heart rate variability analysis. http://www.biopac.com/Curriculum/pdf/h32.pdf. Accessed 1 March 2012.

Bower, G. H., & Forgas, J. P. (2001). Mood and social memory. In J. P. Forgas (Ed.), *Handbook of affect and social cognition* (pp. 95–120). Oxford: Pergamon.

Brown, K. G. (2001). Using computers to deliver training: which employees learn and why? *Personnel Psychology, 54*(2), 271–296. doi:10.1111/j.1744-6570.2001.tb00093.x.
Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (Eds.). (2007). *Handbook of psychophysiology* (3rd ed.). New York: Cambridge University Press.
Campbell, D. T., Stanley, J. C., & Gage, N. L. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin.
Caspi, A., & Blau, I. (2008). Social presence in online discussion groups: testing three conceptions and their relations to perceived learning. *Social Psychology Education, 11*, 323–346.
Clore, G. L., Gasper, K., & Garvin, E. (2001). Affect as information. In J. P. Forgas (Ed.), *Handbook of Affect and Social Cognition* (pp. 121–144). Nahwah: Erlbaum.
Conati, C. (2002). Probabilistic assessment of user's emotions in educational games. *Applied Artificial Intelligence, 16*(7–8), 555–575.
Conati, C., & Zhao, X. (2004). Building and evaluating an intelligent pedagogical agent to improve the effectiveness of an educational game. In 9th International Conference on Intelligent User Interfaces (pp. 6-13). Funchal, Madeira, Portugal: ACM.
Craig, S. D., Graesser, A. C., Sullins, J., & Gholson, B. (2004). Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media, 29*(3), 241–250.
Csíkszentmihályi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.
Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: on the nonconscious operation of behavioral goals. *Journal of Personality and Social Psychology, 89*(2), 129–142.
Cytowic, R. E. (1989). *Synesthesia: A Union of the Senses*. New York: Springer-Verlag.
D'Mello, S. (2008). Automatic detection of learner's affect from conversational cues. *User Modeling and User-Adapted Interaction, 18*(1–2), 45–80. doi:10.1007/s11257-007-9037-6.
D'Mello, S., & Graesser, A. (2012a). AutoTutor and affective autotutor learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems, 2*(4), 1–39. doi:10.1145/2395123.2395128.
D'Mello, S., & Graesser, A. (2012b). Emotions during learning with AutoTutor. In P. Durlach & A. Lesgold (Eds.), *Adaptive Technologies for Training and Education* (pp. 117–139). Cambridge: Cambridge University Press.
D'Mello, S., Lehman, B., & Graesser, A. (2011). A motivationally supportive affect-sensitive AutoTutor. In R. A. Calvo, & S. K. D'Mello (Ed.), New perspectives on affect and learning technologies (Vol. 3, pp. 113-126, Explorations in the Learning Sciences, Instructional Systems and Performance Technologies): Springer New York.
Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace. *Journal of Applied Social Psychology, 22*(14), 1111–1132. doi:10.1111/j.1559-1816.1992.tb00945.x.
Ellis, H. C., & Ashbrook, P. W. (1988). Resource allocation model of the effects of depressed mood states on memory. In K. Fiedler & J. P. Forgas (Eds.), *Affect, Cognition and Social Behavior* (pp. 25–43). Gottingen: Hogrefe.
Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: the processing efficiency theory. *Cognition and Emotion, 6*(6), 409–434. doi:10.1080/02699939208409696.
Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion, 7*(2), 336–353.
Frijda, N. H. (1986). *The Emotions*. Cambridge: Cambridge University Press.
Fu, F., Su, R., & Yu, S. (2009). EGameFlow: a scale to measure learners' enjoyment of e-learning games. *Computers & Education, 52*(1), 101–112. doi:10.1016/j.compedu.2008.07.004.
Ghani, J., & Deshpande, S. P. (1994). Task characteristics and the experience of optimal flow in human computer interaction. *The Journal of Psychology, 128*(4), 381–391. doi:10.1080/00223980.1994.9712742.
Goleman, D. (1995). *Emotional Intelligence*. New York: Bantam Books.
Hernández, Y., Sucar, L. E., & Conati, C. (2008). An affective behavior model for intelligent tutors. In Proceedings of 9th International Conference on Intelligent Tutoring Systems (pp. 819-821). Montreal, Canada: Springer-Verlag.
Hernández, Y., Sucar, L. E., & Conati, C. (2009). Incorporating an affective behavior model into an educational game. In Twenty Second International FLAIRS Conference. Florida, USA.
Hone, K. (2006). Empathic agents to reduce user frustration: the effects of varying agent characteristics. *Interacting with Computers, 18*(2), 227–245.
Hsu, C.-L., & Lu, H.-P. (2004). Why do people play on-line games? an extended TAM with social influences and flow experience. *Information & Management, 41*(7), 853–868. doi:10.1016/j.im.2003.08.014.

Jiang, M. (2000). A study of factors influencing students' perceived learning in a web-based course environment. *International Journal of Educational Telecommunications, 6*(4), 317–338.
Kaklauskas, A., Kuzminske, A., Zavadskas, E. K., Daniunas, A., Kaklauskas, G., Seniut, M., et al. (2015). Affective tutoring system for built environment management. *Computers & Education, 82*, 202–216. doi:10.1016/j.compedu.2014.11.016.
Kim, C. (2012). The role of affective and motivational factors in designing personalized learning environments. *Educational Technology Research and Development, 60*(4), 563–584.
Kim, Y., Baylor, A. L., & Shen, E. (2007). Pedagogical agents as learning companions: the impact of agent emotion and gender. *Journal of Computer Assisted learning, 23*(3), 220–234.
Kirsch, D. (1997). The Sentic Mouse : Developing a tool for measuring emotional valence. http://affect.media.mit.edu/projectpages/archived/projects/sentic_mouse.html. Accessed 5th Nov 2012.
Klein, J., Moon, Y., & Picard, R. W. (2002). This computer responds to user frustration: theory, design and results. *Interacting with Computers, 14*(2), 119–140.
Kort, B., Reilly, R., & Picard, R. W. An affective model of interplay between emotions and learning: Reengineering educational pedagogy—Building a learning companion. In IEEE International Conference on Advanced Learning Technologies, Madison, USA, 2001 (pp. 43-48).
Lee, M. K. O., Cheung, C. M. K., & Chen, Z. (2005). Acceptance of internet-based learning medium: the role of extrinsic and intrinsic motivation. *Information & Management, 42*(8), 1095–1104.
Lin, H. C. K., Chao, C.-J., & Huang, T.-C. (2015). From a perspective on foreign language learning anxiety to develop an affective tutoring system. [journal article]. *Educational Technology Research and Development., 63*(5), 727–747. doi:10.1007/s11423-015-9385-6.
Lin, A. C. H., Fernandez, W. D., & Gregor, S. (2012). Understanding web enjoyment experiences and informal learning: a study in a museum context. *Decision Support Systems, 53*(4), 846–858. doi:10.1016/j.dss.2012.05.020.
Lin, H. C. K., Su, S. H., Chao, C. J., Hsieh, C. Y., & Tsai, S. C. (2016). Construction of multi-mode affective learning system: taking affective design as an example. *Educational Technology & Society, 19*(2), 132–147.
Lisetti, C. L. A user model of emotion-cognition. In Workshop on Attitude, Personality, and Emotions in User-Adapted Interaction at the International Conference on User-Modeling (UM'99), Banff, Canada, 1999.
Litman, D. J., & Silliman, S. (2004). ITSPOKE: An intelligent tutoring spoken dialogue system. In Human Language Technology Conference 4th Meeting of the North American Chapter of the Association for Computational Linguistics (pp. 5-8). Boston, USA: Association for Computational Linguistics.
Lowendahl, J.-M. (2012). Hype Cycle for Education. http://www.gartner.com/DisplayDocument?doc_cd=233974&ref=g_sitelink. Accessed 1st September 2012.
Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 5*(4), 349–361.
Mayer, R. E. (1998). A split-attention effect in multimedia learning: evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312–320. doi:10.1037/0022-0663.90.2.312.
Mayer, R. E. (2005). *Cognitive Theory of Multimedia Learning*. New York: Cambridge University Press.
Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction, 33*, 12–18. doi:10.1016/j.learninstruc.2014.02.004.
Microsoft Corporation (2009). Microsoft Agent. http://www.microsoft.com/products/msagent/main.aspx. Accessed 22 June 2012.
Molster, C., Charles, T., Samanek, A., & O'Leary, P. (2009). Australian study on public knowledge of human genetics and health. *Public Health Genomics, 12*(2), 84–91.
Moreno, R. (2005). Multimedia learning with animated pedagogical agents. In R. Mayer (Ed.), *TheCambridge Handbook of Multimedia Learning* (pp. 507–524). Cambridge: Cambridge University Press.
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*(3), 149–158.
Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.
Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543–578.
Picard, R. W. (1997). *Affective Computing*. Massachusetts: MIT Press.
Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: effects of shape and color on affect and learning. *Learning and Instruction, 29*, 128–140. doi:10.1016/j.learninstruc.2013.02.006.
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 (Ed.),

Tutorial and Research Workshop on Affective Dialogue Systems 2004 (pp. 53-64, Lecture Notes in Computer Science). Kloster Irsee, Germany: Springer Berlin/Heidelberg.
Richards, M. (1996). Lay and professional knowledge of genetics and inheritance. *Public Understanding of Science, 5*(3), 217–230.
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.
Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behaviour* (pp. 527–561). New York: Guildford Press.
Schwarz, N., & Clore, G. L. (1988). How do I feel about it? the informative function of affective states. In K. Fiedler & I. Forgas (Eds.), *Affect, Cognition, and Social Behavior* (pp. 44–62). Göttingen: Hogrefe.
Shen, L., Wang, M., & Shen, R. (2009). Affective e-learning: using "emotional" data to improve learning in pervasive learning environment. *Educational Technology & Society, 12*(2), 176–189.
Sofer, W., & Gribbin, M. (2010). Morgan : A genetics tutorial. http://morgan.rutgers.edu/MorganWebFrames/How_to_use/HTU_Frameset.html. Accessed 1 August 2010.
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: Erlbaum.
Susarla, S., Adcock, A., Van Eck, R., Moreno, K., & Graesser, A. Development and evaluation of a lesson authoring tool for AutoTutor. In Artifical Intelligence in Education Conference, Sydney, Australia, 2003 (pp. 378-387).
Thompson, N., Koziniec, T., & McGill, T. (2012). An open affective computing platform. In Proceedings of the IEEE 3rd International Conference on Networked and Embedded Systems for Every Application (pp. 1-10). Liverpool, UK.
Thompson, N., & McGill, T. (2015). Affective stack—a model for affective computing application development. *Journal of Software, 10*(8), 919–930.
Thought Technology (2010). CardioPro Infiniti HRV analysis module user manual. http://www.thoughttechnology.com/pdf/manuals/SA7590%20CardioPro%20Infiniti%20HRV%20Analysis%20Module%20User%20Manual.pdf. Accessed 1st February 2012.
Um, E., Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). Emotional design in multimedia learning. *Journal of Educational Psychology, 104*(2), 485–498.
van der Meij, H. (2013). Motivating agents in software tutorials. *Computers in Human Behavior, 29*(3), 845–857. doi:10.1016/j.chb.2012.10.018.
Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76*(2), 92.
Woolf, B., Burelson, W., & Arroyo, I. Emotional intelligence for computer tutors. In AIED (Ed.), 13th International Conference on Artificial Intelligence in Education, Los Angeles, USA, 2007 (pp. 6-15).
Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: recognising and responding to student affect. *International Journal of Learning Technology, 4*(3), 129–164.
Wu, C.-H., Huang, Y.-M., & Hwang, J.-P. (2015). Review of affective computing in education/learning: trends and challenges. *British Journal of Educational Technology,*. doi:10.1111/bjet.12324.
Yannakakis, G., Hallam, J., & Lund, H. (2008). Entertainment capture through heart rate activity in physical interactive playgrounds. *User Modeling and User-Adapted Interaction, 18*(1), 207–243. doi:10.1007/s11257-007-9036-7.
Zakharov, K., Mitrovic, A., & Johnston, L. (2007). Pedagogical agents trying on a caring mentor role. *Frontiers in Artificial Intelligence and Applications, 158*, 59–66.

Nik Thompson is a Senior Lecturer in Information Systems at Curtin University. He has a Ph.D. from Murdoch University in Western Australia in the area of Affective Computing. His research interests include e-learning, physiological computing and the development of adaptive computer interfaces that detect and respond to the emotional state of the user.

Tanya Jane McGill is an Associate Professor in Information Technology at Murdoch University in Western Australia. She has a Ph.D. from Murdoch University. Her major research interests include e-learning, ICT education, and information security. Her work has appeared in various journals including Computers & Education, Decision Support Systems, Journal of Computer Assisted Learning, Behaviour and Information Technology and Journal of Organizational and End User Computing.