PlayPhysics:
An Emotional Student Model for Game-based Learning

Karla Cristina Muñoz Esquivel
B.Sc. (Hons.), M.Sc., P.G. Cert., CAPM & AFHEA

School of Computing & Intelligent Systems
Faculty of Computing & Engineering
University of Ulster, Magee

A thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

July, 2012

Table of Contents

List of Figures

List of Tables

List of Acronyms

AI	Artificial Intelligence
ANNs	Artificial Neural Networks
API	Application Programming Interface
BBNs	Bayesian Belief Networks
BDI	Belief-Desire-Intention
BPMS	Body Posture Measurement System
CAI	Computer-Assisted Instruction
CASE	Computer-Aided Software Engineering
CB-AUM	Cognitive-Based Affective User Modelling
CCM	Campus Ciudad de México
CIDRs	Conditional Independent or Dependent Relations
CPTs	Conditional Probability Tables
CSS	Cascading Style Sheets
DBNs	Dynamic Bayesian Networks
DGD1	Demographic Game Design model
ECA	Embodied Conversational Agent
EDA	Electro Dermal Activity
EDM	Educational Data Mining
EGEMS	Electronic Games for Education in Mathematics and Science
EM	Expectation Maximisation
EPA	Embodied Pedagogical Agent
FIDGE	Fuzzified Instructional Design Development of Game-like Environments
FSM	Finite State Machine
GBL	Game-based Learning
GSR	Galvanic Skin Response
GUI	Graphical-User Interface
HCI	Human Computer Interaction
HR	Heart Rate
HTML	HyperText Markup Language
IA	Intelligent Agent
IDDMs	Instructional Design and Development Models
IT	Information Technology
ITESM-CCM	Tecnológico de Monterrey, Mexico City Campus
ITSs	Intelligent Tutoring Systems
JSPs	Java Server Pages
LCP	LEGO Communication Protocol
LeJOS	LEGO Java Operating System
MIT	Massachusetts Institute of Technology
MLR	Multinomial Logistic Regression
MVC	Model-View-Controller
NPCs	Non-Player Characters
NXT	NeXT generation LEGO

OLEs	Open-ended Learning Environments
OLMs	Open Learner Models
OMG	Object Management Group
OO	Object Oriented
PC	Peter-Clarkson
PRMs	Probabilistic Relational Models
SAL	Sensitive Artificial Listener
SEM	Structural Equation Modelling
SI	International System of Units (‘Système international d’unités’)
SPSS	Statistical Package for the Social Sciences
SVM	Supported Vector Machine
UML	Unified Modelling Language
URL	Uniform Resource Locator
VLEs	Virtual Learning Environments
XML	eXtensible Markup Language

Acknowledgments

First, I would like to convey my gratitude to my supervisors, Prof. Paul Mc Kevitt and Dr. Tom Lunney, who have guided, advised and supported me all the way through this research. I would like to express my appreciation to our external collaborators, Dr. Julieta Noguez and Dr. Luis Neri at Tecnológico de Monterrey, Mexico City campus (ITESM-CCM), for sharing their expertise, guidance and suggestions to this research.

During this research experience, feedback from colleagues has been key to my progress, therefore I want to convey my thanks to the reviewers of the 100 Day review, the 1st year report and the 3rd year presentation, namely Dr. Joan Condell, Dr. N.H. Siddique, Prof. Sally McClean and Head of Research Graduate School, Dr. Philip Morrow. I would like to thank my colleagues in the Intelligent Systems Research Centre (ISRC) for their friendship and technical advice, particularly, Dr. Girijesh Prasad, Dr. Damien Coyle, Dr. Abdul Satti, Gabriel Adrian Deak, Ted Leath, Pat Kinsella, Bernard McGarry, Neil McDonnell, Paddy McDonough and Pat Conlon. I would like to thank Lee Cadieux from the Art and Design Research Institute for his advice on animation techniques and 3D modelling and Prof. Mark Shevlin from the Psychology Research Institute, for his guidance regarding statistical analysis. I wish to acknowledge colleagues from the School of Psychology, Queen's University Belfast, Dr. Deaglan Page and Dr. Donncha Hanna who have given advice on statistical textbooks and published methodologies.

Also, I wish to acknowledge the support given by our external collaborators at ITESM-CCM and the other members of the *e-learning* research team, Victor Robledo, Dr. Moises Alancastre, Dr. Lourdes Muñoz, Dr. Gilberto Huesca, Gerardo Aguilar and Benjamín Hernández, without them, the testing of *PlayPhysics* just would not have been possible. Thanks to all the participants of this research, especially high school and undergraduate students of Engineering at ITESM-CCM from 2010 to 2012, who gave us constructive feedback and helped us to evaluate our emotional student model and the *PlayPhysics platform*.

Thanks to Richard Walsh from *ZooCreative* for his assistance in the 3D modelling of *PlayPhysics*' player characters and Peter Starostin whose free rig, *LowMax,* was adapted to create PlayPhysics' learning companion, the *M8* robot.

I would like to convey my gratitude to the University of Ulster for supporting me for three years with a Vice Chancellor's Research Scholarship (VCRS) and to the European Union Social Signal Processing Network (SSPnet) & the Scottish Informatics and Computer Science Alliance (SICSA) for the scholarship granted to attend the First Summer School of Affective Computing and Social Signal Processing (ACSSP10) at the University of Edinburgh.

In addition, I would like to thank the constructive comments and feedback provided by reviewers in international conferences, workshops, journals and books. In particular, I would like to thank the reviewers of Human-Machine Interaction Network of Emotion (HUMAINE).

I would like to finish by thanking the most important people in my life, my family. Brad Henry states:

> *"Families are the compass that guides us. They are the inspiration to reach great heights, and our comfort when we occasionally falter"*
>
> (Johnson 2010, p. 61)

I am very grateful to my parents Juan Carlos and María Josefa, my brother Juan Carlos and my grannies Marina and Amalia, who taught me to fight for what I want and to challenge myself everyday and who always have been by my side regardless of distance. Particularly, I want to thank my beloved husband, Dennis Heaney. Thank you for believing in me, for your patience, for sharing your technical expertise and your passion for knowledge. You are my favourite person in this world and without you I would not see this work reach completion today.

Abstract

Game-based learning (GBL) environments introduced a new generation of Intelligent Tutoring Systems (ITSs) that provide personalised instruction by being constantly aware of student reactions to the system. Student motivation, attitudes, self-efficacy and affective state have been the key focus of such developments. Current models of student emotion have shown promise in laboratory environments. However, the problem of accurately recognising and inferring student emotions within learning environments persists. The majority of already existing computational models of student emotion employ cognitive theories that are not derived from the learning context.

Control-value theory (Pekrun et al. 2007) assumes that control and value appraisals are the most meaningful for determining emotions in educational settings. Our proposed computational emotional student model uses the Control-value theory for reasoning about learners' emotions in GBL environments settings. The main hypothesis is that this model will recognise student *achievement* emotions, i.e. emotions relevant to the educational context, with reasonable accuracy (not random). The definition, implementation and evaluation of our computational emotional student model in PlayPhysics, an emotional game-based learning environment for teaching physics, are discussed. Our emotional model is implemented with a dynamic sequence of Bayesian networks for representation of learners' *achievement emotions*. Probabilistic Relational Models (PRMs) are employed to facilitate their derivation. The Necessary Path Condition algorithm is employed in combination with Pearson correlations and Binary and Multinomial logistic regression for defining network structure. The Expectation Maximisation (EM) learning algorithm is employed for network parameter learning. Our model employs answers to questions in-game dialogues, contextual variables and physiological variables for recognising student emotion.

Results show a fair accuracy of classification of student achievement emotions for the PlayPhysics' emotional student model when only contextual and behaviour variables are considered (values of Cohen's Kappa in a range larger than 0.2 but lower than or equal to 0.4), which then improves when physiological variables, i.e. Galvanic Skin Response (GSR), are incorporated (values of Cohen's Kappa in a range larger than 0.4 and lower than or equal to 0.6). Our emotional model provides enhanced understanding about the factors involved in reasoning about emotion. PlayPhysics GBL environment is assessed to attain an enhanced understanding of the student experience of *achievement emotions*. Future work may focus on creating further game challenges, identifying enhanced predictors for *control* and *value*, e.g. using sentiment analysis and analysis of facial expressions. Numerous applications, in areas ranging from biology to e-commerce, are envisioned for the application of our approach to create intelligible and dynamic genetic and emotional consumer data models.

Note on Access to Contents

I hereby declare that with effect from the date on which the thesis is deposited in the Library of the University of Ulster, I permit

1. The Librarian of the University to allow the thesis to be copied in whole or in part without reference to me on the understanding that such authority applies to the provision of single copies made for study purposes or for inclusion within the stock of another library.
2. The thesis to be made available through the Ulster Institutional Repository and/or EthOS under the terms of the Ulster eThesis Deposit Agreement which I have signed.

Karla Muñoz

Chapter 1: Introduction

Nowadays students grow up in an environment highly influenced by information technology (IT), e.g. the Internet, social media and networking sites, blogs, online video streaming, video game consoles and mobile devices. Digital media is becoming increasingly important in our daily lives, so it is imperative that education adapts to these changes (Oblinger 2004, Woolf 2009). As a result, students have different expectations about how, where and when learning should happen (Oblinger 2004).

Lecturers have always been challenged to fulfil students' expectations with respect to learning and maintaining engagement and the influence of digital technology on students has made the attainment of this goal increasingly complex. In addition, students should not be limited to receiving and retrieving information. They must be capable of applying knowledge in an innovative manner (The ICE House Project Group 2009).

One-to-one human tutoring has proven the most effective teaching and learning strategy (Bloom 1984), e.g. enabling students to attain the highest scores and decreasing the time spent on learning to a third of the time spent in group sessions. However, Intelligent Tutoring Systems (ITSs) have shown the same effectiveness and decreased time spent on learning to one-half (Regian et al. 1996). ITSs have been characterised as software incorporating artificial intelligence (AI) techniques (Woolf 2009) in order to achieve two main objectives: (1) reasoning about students' features and needs, i.e. student modelling, and (2) adapting teaching strategies, i.e. tutor modelling.

Online learning, a form of computer instruction, enables students to achieve non-traditional, flexible and interactive learning through the use of technology (Moore et al. 2011). It is an enhanced version of *distance learning* that describes education in which the instructor is located in a different physical place from the learner, or where education happens at different times. A Virtual Learning Environment (VLE) facilitates and supports *online learning in the form of simulations* in a web-based area, where students can explore safely, experience and understand the effects of specific phenomena or operation of high cost equipment without the risk of damaging it (Reilly 2008). Game-based learning (GBL) environments, i.e. edutainment, constitute another approach for facilitating online-learning, which has been proven to achieve and maintain student engagement more easily than VLEs due to its additional features, such as the provision of high interactivity, immediate feedback, storytelling and power-ups (game items that bestow the player character with special abilities), which facilitate the

establishment of an emotional link between the student and the game (Sykes 2006). However, a key problem, which influences lecturers' reticence on integrating GBL environments into the syllabus, is the mismatch between curriculum and game content (Carpenter and Windsor 2006) and the fact that learning to play successfully does not ensure achieving learning goals (Conati 2002). To guarantee learning, GBL environments follow design principles and incorporate ITSs.

Computer tutoring, in a similar manner to human tutoring, has focused on knowing how knowledge and understanding is attained by students (Muñoz et al. 2009b, Noguez and Sucar 2005) and what features are identified as the most significant to attain effective tutoring (Lepper et al. 1993, Pearson et al. 2003). It is important to note that effective and expert human tutors are capable of identifying accurately student errors and misconceptions and take into consideration students' personal differences, learning preferences, attitudes, emotional and motivational states simultaneously in order to provide pedagogical feedback. However, it is still not clear, to what degree computer tutoring should resemble human tutoring, the *plausibility problem*, or how capable it would be of conveying appropriate human behaviour (Du Boulay and Luckin 2001, Lepper et al. 1993).

Human tutors consider different kinds of information in order to be aware of student learning and motivational needs, such as written solutions to a problem, body language, voice inflection and intonation, facial gestures, previous social interaction, personal knowledge, context, nature of interactions and activities performed (Alexander and Sarrafzadeh 2008). Therefore, latent and key questions in the field of ITSs are which of these behaviour features should or could be selected and how should this be conceptualised to attain an accurate perception of the student needs. Also, if such conceptualisation and selection of features is attained, how can the accuracy of computer tutor's reasoning be measured. The ultimate goal is to create *positive learning experiences* in student minds, which recalls the question of what defines a positive experience (Westerinck et al. 2008a).

One of the key elements of the learning experience is emotion (Westerinck et al. 2008a, Brave and Nass 2008), which has been shown to influence not only student understanding, learning, motivation and cognitive performance, but also lecturer performance. Teachers in the United States and European have resigned from teaching in the early stages of their careers owing to emotions such as anger and anxiety (Pekrun et al. 2007). In addition, from the social and Human-Computer Interaction (HCI) perspectives, a Graphical-User-Interface (GUI), an educational application, incapable of properly showing emotion, limits performance and is perceived as socially-disabled, useless and untrustworthy (Brave and Nass 2008). This perception is explained by recalling that showing emotion is one of the main building blocks in the creation of social relations (Keltner and Lerner 2010). However, in order to respond appropriately to emotion, it is necessary to first identify it. As a result, the significance

of emotion in education has gained attention, specifically in the fields of psychology and computing tutoring.

Human actions, our own or someone else's, are highly associated with emotions. Not only do emotions stimulate action, but actions also lead to the occurrence of events, which are sometimes considered individually significant, e.g. needs, goals and interests. If the events are relevant to the person, emotion arises as a response (Brave and Nass 2008). Latent questions in this field are: (1) how emotion originates and (2) which emotions are relevant to the specific context, e.g. teaching-learning experience. Additionally, an associated problem is how to respond appropriately to emotion.

This thesis supports the idea that ITSs will be capable of reasoning about student emotions in order to encourage performance and to suitably adapt learning in order to create engaging and meaningful experiences. It is anticipated that ITSs should be able to attain at least the accuracy of humans when recognising emotion, i.e. approximately 70% (Keltner and Lerner 2010).

1.1 Research problems and motivation

Affective Computing is a research area focused on enabling computers to recognise and show emotion (Picard 1995). Key motivations for enabling computers to react to student emotions are: providing suitable support to improve user experiences, facilitating improved performance and encouraging the creation of meaningful relationships with users by promoting user trust and giving a sensation of competence (Brave and Nass 2008). Affective Computing has influenced research in other fields, such as ITSs and serious gaming, providing a fresh outlook for applications. As a result, ITSs and serious games are focused on understanding and expressing emotion (Picard 1995, Sykes 2006). As mentioned earlier, educational games or GBL environments incorporate ITSs to ensure personalised instruction, i.e. awareness of student characteristics and needs, and the achievement of learning goals. Focusing on emotion recognition, an immediate question is what approaches are currently employed by ITSs to achieve this goal.

There are three approaches employed by ITSs for undertaking the detection of emotion: (1) identifying the physical effects of emotion (D'Mello et al. 2008, Sarrafzadeh et al. 2008), (2) reasoning about emotion from its origin (Jaques and Vicari 2007) and (3) a hybrid approach, which comprises a combination of both (Conati and Maclaren 2009). These approaches have various advantages and disadvantages due to their nature, such as not being easily affordable or accessible, prone to hardware failure, not using a theory of emotion focused on the teaching-learning experience or unsuitability for online learning or on-site classroom instruction. It is noted that some approaches have shown more promise than others. However, it is evident that further research must be conducted to enhance the accuracy of

identification of the relevant emotions associated with the educational context. In addition, it is necessary to facilitate the derivation of emotional student models to enable ITSs to progress in order to be more feasible and accessible. Furthermore, to achieve these goals, it is noted that the constructs involved in determining an emotion in the educational context should be more clearly understood.

Simultaneously, research has focused on identifying and adapting learning to student attitudes (Arroyo and Woolf 2005), motivation (Del Soldato and Du Boulay 1995, Rebolledo-Mendez et al. 2006) and self-efficacy (McQuiggan et al. 2008). In these three situations the second approach, which involves reasoning about observable behaviour from its origin, has proven to be more successful than when just used for reasoning about emotion. Also, this approach is suited to supporting *online learning*. Examples of the observable features employed include the time spent on achieving the learning goals, independence and level of performance.

Research has shown that emotion is deeply interrelated with student performance and motivation (Pekrun et al. 2007), which suggests that some of the constructs implicated in determining student levels of motivation or performance are also involved in determining emotion. The questions involved are how these constructs may be related or conceptualised, which of these features should be selected and how the contribution of these features, to classify emotion, can be determined.

In this thesis we focus on *reasoning about emotion from its origin*, since it can be applied during *online learning*. *Online learning* offers the advantage of reaching and assisting a larger number of students anytime–anywhere (Moore et al. 2011), which enables the acquisition of more information about student interaction and enables monitoring of emotions in an environment where learning is more likely to happen rather than a laboratory setting. More specifically, an online GBL environment, which features facilitate the establishment of an emotional link with the student (Sykes 2006). Also, students feel more motivated interacting with GBL environments than when interacting with VLEs (Muñoz et al. 2009b). However, it is important not to overlook the problems involved in Edutainment systems and not to negatively bias user experiences and still achieve the goals for which the application was originally created. Design principles and instructional design development models of Edutainment must be considered for designing and implementing GBL environments. Hence, an evaluation of any system's usability is also important in order to ensure the attainment of the design goals, but also to attain an enhanced understanding of student emotions.

Using *online learning* for supporting the teaching of physics at undergraduate level is also of concern in this thesis, since it is a major challenge to engage students owing to the perceived complexity of physics (Er and Dag 2009) and some students find its underlying structure difficult to understand (Muñoz et al. 2009b). In addition, experts in education in the UK

agreed that education can be improved if the Government effort is focused on providing an enhanced grounding in Science, Technology, Engineering and Mathematics (STEM) (Khan 2011).

Additionally, it is clear that if an approach to *reasoning about observable behaviour from its origin* is employed, a psychological theory or framework, which explains how emotion is originated in the learning context, is more appropriate, since it facilitates the process of derivation of the emotional student model by focusing directly on the relevant emotions. Also, it is more likely that this theory will consider the constructs involved in student motivation and performance. However, currently the majority of research does not focus on identifying emotions related to the learning context or using a theory of emotion that explains emotions in education. The Control-Value theory of achievement emotions (Pekrun et al. 2007) is selected here to derive an emotional student model, since there is currently no computational model of student emotions employing this theory. In addition, Data Mining issues, such as unknown or missing data, also need to be considered, since interaction data for deriving and evaluating the emotional student model can be unintentionally contaminated (Witten and Frank 2005).

Another important question involved in the creation of an emotional student model, is which AI technique is most suited to obtaining an accurate representation of the affective domain given its inherent uncertainty (Witten and Frank 2005). In this thesis a dynamic sequence of Bayesian networks is the preferred technique for representing the domain, it was observed that they enable the use of previous information from the respective domain and the effective handling of uncertainty (Jensen and Nielsen 2007). Additionally, devising an approach for facilitating the derivation of the dynamic sequence of Bayesian networks is also necessary. Hence, a Probabilistic Relational Models (PRMs) approach (Sucar and Noguez 2008) is employed in combination with the application of the Necessary-Path Condition algorithm using Pearson correlations and Binary or Multinomial Logistic Regression (MLR) to solve uncertain associations and define the network structure. Binary or MLR allows the analysis and selection of features by determining not just the significance of the entire model, but also the contribution of the features to the classification of emotion.

Personally, my motivation for this research arose from previous work conducted in my Master's degree which focused on discovering whether game-based learning environments were more effective than virtual learning environments achieving student motivation. This previous work considered the different modalities that educational games can employ to give immediate feedback to student actions (Muñoz et al. 2009b). It highlighted that including an affective dimension in the student model of an ITS may provide more insight into student engagement and performance.

1.2 Aims and objectives

This thesis investigates an original approach to emotional student modelling. The core aim of this work is to advance research in this area through the creation of a computational lmodel of learner emotion, which can accurately represent and recognise relevant student emotions, i.e. *achievement emotions*, using principally contextual, feasible observable variables of low bandwidth related to student behaviour. This model will support mainly online educational gaming, i.e. game-based learning environments. It differs from others in the literature in that it uses Control-Value theory of achievement emotions (Pekrun et al. 2007) as a basis. It is anticipated that our proposed emotional student model will result in a reasonable accuracy in recognising, representing and inferring emotions. Ideally, the model should achieve the precision of an expert individual identifying emotion (Keltner and Lerner 2010). The main objectives of this research are:

- Determining whether answers to questions in game dialogues and observable variables corresponding to student interaction and behaviour can be effective predictors of student emotion in a GBL environment.
- Develop a computational model of student emotions employing Control-value theory (Pekrun et al. 2007), which can be applied to GBL environments. This emotional model comprises a dynamic sequence of Bayesian Belief Networks (BBNs) for emotion representation together with contextual (e.g. mouse focus and requests for help) or physiological variables, i.e. Galvanic Skin Response (GSR), for emotion recognition.
- Test the following hypothesis: the computational model will achieve a reasonable accuracy of classification of student emotions in GBL environment settings, i.e. not random (a value of Cohen's Kappa equal to zero; $\kappa=0$).
- Design, implement and evaluate PlayPhysics, an emotional game-based learning environment for teaching physics, which implements our emotional student model.
- Gain more insight into student experience of achievement emotions in GBL environments, which can assist in the identification of other factors that are influencing student emotion and can be employed in further work to enhance the accuracy of the proposed emotional student model.
- Investigate if any variables of high-bandwidth, which may be made available by the utilisation of sensors - e.g. Galvanic Skin Response (GSR) sensor - can enhance emotion recognition accuracy of the student model.

1.3 Thesis outline

This thesis is comprised of eight chapters. Chapter 2 focuses on discussing work on Emotion from Neuropsychology, Cognitive Psychology and Biology perspectives. It reviews research that attempts to clarify which constructs are relevant in determining emotion, with an emphasis on theories and frameworks that endeavour to explain the origin of emotion in a teaching-learning context. Chapter 3 focuses on investigating the fields of adaptable computer tutoring, i.e. ITSs and Edutainment, discussing their advantages and issues, but also in the case of Edutainment, indicating its relation to emotional aspects that facilitate learning and make it meaningful. The chapter then focuses on Affective Computing and its influence in both research fields, ITSs and Edutainment, in order to set the context for the current approaches employed by ITSs to recognise emotion, i.e. affective student modelling, and their applications.

Chapter 4 covers the generic design of our proposed computational emotional student model using Control-value theory (Pekrun et al. 2007) for reasoning about student emotion, in addition to describing the approach employed in this dissertation for modelling the affective domain and knowing how the proposed affective student model performs over fresh data. A dynamic sequence of BBNs is explored in order to achieve an intelligible and accurate representation and recognition of student emotions. Contextual variables (e.g. mouse focus and requests for help), answers to questions in game dialogues and physiological signals are employed for emotion recognition. Techniques that facilitate the derivation of Bayesian networks, such as PRMs, structural and parameter learning algorithms are employed. We test the assumption that the Bayesian model skeleton can be defined using a combination of the Necessary-Path Condition algorithm, Pearson correlations and the result of the Binary or Multinomial Logistics Regression (BLR/MLR), since Bayesian models are a sub-class of logistic regression and logistic regression provides the advantage of defining the contribution of each regressor to the overall classification. As a result, these two techniques complement each other, since the Necessary-Path Condition algorithm requires that uncertain relations are solved by the domain expert. Finally, we propose to employ cross-validation to estimate the reasoning performance of the emotional student model over fresh data.

In Chapter 5, the design of PlayPhysics, an emotional game-based learning environment for teaching physics, is described. PlayPhysics implements our computational model of student emotions. This chapter discusses the analysis of lecturer and student requirements for giving and receiving education in an introductory course of physics at the undergraduate level. PlayPhysics' functional requirements are presented using Unified Modelling Language (UML). Olympia architecture, combining ITSs and GBL environments, is comprised of several modules in order to provide adaptable and personalised instruction. Specific modules of

Olympia, the input and behaviour analysis modules, are involved in the acquisition of data to completely define and validate the proposed emotional student model. The instantiation of our proposed emotional student model to the specific case study of PlayPhysics is illustrated. The Physics domain used as a reference for designing PlayPhysics game challenge is discussed in addition to the design of a method to identify and respond to student misconceptions. A functional description of PlayPhysics' GUIs is also included.

Chapter 6 discusses the implementation of PlayPhysics. This chapter shows the manner in which Olympia's modules were adapted to develop the PlayPhysics prototype. The problems that are addressed are how to collect student interaction data, i.e. contextual variables (e.g. mouse focus and requests for help), biofeedback signals and how to enable student self-reporting. This interaction data is employed to define and evaluate our proposed emotional student model in Chapter 7. PlayPhysics self-reporting capabilities are supported through an EmoReport wheel, which can be used by students at anytime to self-report their emotional state. The operation of *M8-robot*, PlayPhysics' learning companion is used as an exemplar. A version of PlayPhysics was created to support the recording of the student Galvanic Skin Response (GSR) signal through a Bluetooth biofeedback device. The latter was created using the LEGO NXT intelligent brick and LEJOS. PlayPhysics was implemented using Java, the Unity 3D game engine, JavaScript and other highly useful software tools that speeded up the process. Screen-shots exemplifying user interaction are presented.

Chapter 7 discusses the experimental results of testing our computational emotional student model and this thesis hypothesis, i.e. the model will achieve a reasonable accuracy of classification of student emotions in GBL environments (not random). Ten-fold stratified cross-validation is employed to assess the performance of the PlayPhysics' emotional student model over fresh data. Cohen's Kappa is employed to evaluate the agreement between student self-reports (observed values) and the predictions achieved by PlayPhysics (predicted values). The approach and results show promise. PlayPhysics' emotional student model, using only contextual-variables, is capable of achieving a fair accuracy (values for Cohen's Kappa in a range between 0.2 and 0.4) of classification of achievement emotions. However, the model could potentially be refined, so future work is required to find other factors that can enhance the accuracy of classification of *control* and *value*. It was observed that the accuracy of PlayPhysics' emotional student model increases when using a combination of contextual variables and physiological signals. In addition, with a view to providing more insight about student experience of achievement emotions in PlayPhysics' GBL environment, the results of a qualitative and quantitative evaluation are presented.

Chapter 8 concludes with a summary of the thesis, its relation to other work and prospects for future work. Our model was derived using Control-value theory (Pekrun et al. 2007), a Cognitive-Based Affective User Modelling approach (CB-AUM), a dynamic sequence of

Bayesian networks for representation and contextual or physiological variables (GSR) for reasoning about student achievement emotions. The resultant model assists us in achieving an enhanced understanding of our student participants. We employed percentages to analyse the accuracy of our emotional model, but we also employ values of Cohen's Kappa to demonstrate that the agreement is not random, which has not been explored in related work. The contextual variables employed in our model have been used previously to recognise student motivation and self-efficacy. Here we show that these variables can also be employed for reasoning about student emotion. We conclude that our computational model shows promise, but also presents opportunity for enhancement. Future work will focus on identifying other random variables that may improve the classification of control and value, e.g. facial gestures, speech intonation and sentiment analysis. Useful applications of our approach for creating other inteligible and temporal data models related to contexts such as biology and e-Commerce are highlighted.

Chapter 2: Emotion and Education

Emotion is understood from Biological, Neuropsychological and Cognitive psychological viewpoints, its relevance in human experiences, in human-computer interaction (HCI) and in the educational contexts.

Emotion is important because it is an inherent part of human experience. It can encourage or discourage action, gives meaning to and enhances experience (Brave and Nass 2008, Westerinck 2008). On the other hand, interacting with computers has also become a natural part of people's lives. As a result, emotions or affective states also play a critical role in every goal, activity or task related to the use of computers, e.g. performing calculations in an Excel spreadsheet, creating a 3D model in Autodesk Maya and performing an online search in a browser. Research has shown that Graphical User Interfaces (GUIs) that do not consider students' affective state or do not convey a suitable affective response may limit students' performance (Brave and Nass 2008). Emotion is comprised of behavioural, psychological, conceptual and cognitive components (Ortony et al. 1990).

2.1 Understanding emotion

From a neuropsychology viewpoint, a model explaining how the human brain processes information and how emotion arises is the model by LeDoux (1996). According to this model, there are three key parts of the brain involved in the experience of emotion: thalamus, cortex and limbic system (see Figure 2.1). The thalamus receives all the information collected through the senses from the environment and processes it, i.e. it is like a basic signal processor. Afterwards and in parallel, information is sent to the cortex for further or high level processing and also to the limbic system. The limbic system is where emotion is generated, since its function comprises evaluating the stimulus or inputs and assessing their relevance against the individuals' needs and goals. If relevance exists, the limbic system sends information to the body in the form of signals in order to control physiological responses and to the cortex in order to influence attention and other cognitive processes.

From the basic processing of signals originated by the communication between the thalamus and the limbic system, primary emotions or an unconscious experience of emotion arises such as feeling suddenly frightened, feeling an instinctive aversion or feeling an instinctive attraction. Samples of events related to work with computers, which may influence

the experience of primary emotions, according to Brave and Nass (2008) are: animations that appear suddenly, the opening of pop-up windows or a Non-Player Character (NPC) walking towards the user. Secondary emotions, e.g. pride and frustration, occur once that high-level processing happens, which requires employing as reference a knowledge-base. The connection between the limbic system and the cortex is the main source of secondary emotions. Both connections, thalamic-limbic and limbic-cortex cause the conscious experience of emotions and perhaps other emotions. As an example Brave and Nass (2008) mention that someone who initially was frightened due to a stimulus, for example by accidentally deleting a file. This initial emotion may be reinforced or weakened through rational evaluation and as a result this person may experience other emotions. If the user discovers that the file was not deleted, but was filed under another folder, the user will feel relief. However, if the user confirms that the file was deleted and it is not possible to recover it, the user may feel hopeless. The conscious experience of emotion takes place when the cortex receives feedback from the limbic system and the body (LeDoux 1996).

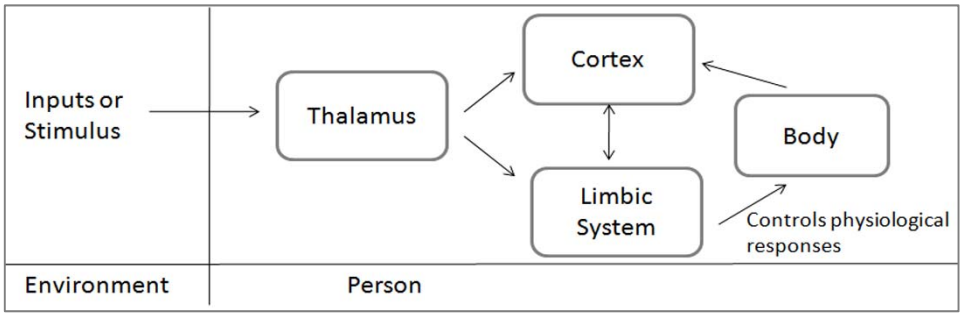


Figure 2.1 Graphic representation of LeDoux's model of emotion

From the investigation of expressions in animals and man by Darwin (1872), it is believed that human emotions are inherited. Therefore, emotions were acquired for preservation in order to adapt to the surrounding environment. In 1868, Darwin conducted an experiment where he showed photographs of a face that were originally taken by Guillaume-Benjamin Duchenne, a French anatomist, to twenty-three guests at dinner parties and asked them to identify the emotion expressed by the face. The photos were of a man whose face was paralysed. Electrodes were used to shape his face into different expressions. Darwin's findings demonstrated that the majority of his participants' assessments agreed on specific photographs related to emotion, such as fear, surprise, happines, sadness and anger (Pearn 2011, Hegarty 2011). In 2011, this experiment was recreated by the Darwin Correspondence Project Team (2012) at Cambridge, which called for participants worldwide to label the same eleven photographs. As an additional experiment they asked participants to label emotions in videos. The team collected in total eighteen thousand responses to the original photograph labelling task. Their main objectives were to gain more knowledge of the context in which

Darwin's experiment took place and the problems he faced. Their findings show that responses were diverse. There was more than 50% agreement in photographs 2, 3 and 4, which show 'surprise', 'terror' and 'grief & despair' respectively, and over 25% associated photograph 1 – which showed 'laughter'. These findings agreed with Darwin's own conclusions, which state that some of these artificially created facial expressions convey emotion more convincingly than others. The second experiment, entitled 'on-line video annotation', is ongoing and led by Peter Robinson. This experiment is implemented using web-based tools with the main objective of creating a bank of facial expressions and emotions, which can be used to create intelligent interfaces with applications in satellite navigation and teaching (Robinson and El Kaliouby 2009).

According to Darwin's perspective, emotions are culturally dependent. However, from the investigation by Ortony et al. (1990) entitled 'the Cognitive Structure of Emotions', it is noted that it is also believed that some emotions are not inherited and are the result of social interactions and constructs and require high level cognitive processes that involve a knowledge-base. As a result, it is observed that according to this theory, emotions may not depend on individuals' culture. The Ortony, Clore and Collins (OCC) model of emotion explains that different types of emotions arise depending on the elements of cognition that participate in the appraisal (Ortony et al. 1990). According to the OCC theory individuals can focus on agents, events or objects. In order to focus on one of these elements, it should be an individual's social standards, goals or attitudes against which the actions of agents, the consequences of events and characteristics of objects are evaluated. For example, when we are pleased with the consequences of an event and the event is relevant to our own goals according to the OCC model, it is very likely that we feel satisfaction, which is considered a prospect-based emotion. The derivation of this taxonomy and its evaluation was achieved through a case study, which involved keeping a personal diary. In addition, it is important to note that self-reports of emotions are not objective or conclusive. However, self-reports are taken as valid, because only people having the emotion have access to it, they are the result of subjective experiences, e.g. like colour or pain. The OCC theory has been employed frequently to derive computational and affective student models.

Damasio (2000) and Darwin (1872), mention that non-human creatures are capable of experimenting and expressing emotions. However, in humans, emotions are associated with complex ideas such as values, judgements and principles, which is the key reason for the importance of emotions. According to Damasio (2000) emotion engenders feelings corresponding to a biological process. Feelings are inwardly directed and private, as opposed to emotions, which are outwardly directed and public. However, Damasio (2000) emphasises that the duration and impact of feelings depends on a person's level of consciousness, i.e.

knowing that a feeling is being experienced. Once a person knows that he/she has a feeling, Damasio (2000) suggests that this feeling may be represented as neural or mental patterns.

Damasio (2000) stresses the relationship between emotions and homeostasis. The latter is related to the regulation and automation of physiological reactions that stabilise the internal states of a living organism, e.g. body temperature and quantity of oxygen in our blood. Also, Damasio (2000) demonstrated the contribution of emotion to decision making and reasoning processes. The experiment involved participants that were victims of neurological damage, which affected brain areas involved in experiencing specific emotions.

Wittgenstein's (1953/2009) approach focuses on achieving a summarised perspective of everyday psychological words/concepts and the relationships between them, imagination and the external-world. He considers what makes sense when talking about emotion instead of defining it. Wittgenstein argues that emotions colour thoughts, i.e. thoughts are classified in categories with characteristics, such as joy, and are part of the subset of concepts of experience or sense-perceptions, to which verbs and sensations also belong. The latter have a duration and degree, are interconnected, can be qualitatively mixed and serve the purpose of transferring information from the real-world.

However, there are also theories that believe that there are a set of basic or universal emotions between individuals as argued by Ekman (1999), who said that there is a set of emotions, e.g. anger, contempt, sadness, disgust, fear and surprise, which do not have a cultural origin, are innate and are characterised by specific facial gestures. The judgement of these emotions through facial expressions surpasses chance, e.g. achieving accuracies between 60% and 80%.This theory has also been employed on several occasions for computational and affective student modelling. Consequently, Keltner and Ekman (2000) endeavour to measure facial activity through facial changes, instead of employing electromyography, which is less precise and more intrusive. Also, Keltner and Ekman (2000) emphasise the debate that exists about visualising emotions as separate or interrelated entities. The evolutionary view considers that each emotion serves a specific adaptive function that is associated with specific responses. On the other hand, the dimensional viewpoint sees emotions as inseparable, but their measurement can be enhanced if visualised as dissimilar in degree of dimension, e.g. valence, activity, approach or withdrawal. The dimensional viewpoint is preferred by those who consider emotion as being socially learned or culturally variable.

Keltner and Ekman (2000) argue that diverse regions of the nervous system are employed in order to express and perceive different emotions. Therefore, whilst analysing the meaning of facial expressions, supporters of the dimensional view state that first it is necessary to focus on determining the valence of the facial expression, since it is believed that the same region of the brain is employed to identify emotions from the same valence, e.g. negative or positive. Those who support the dimensional view focus on the level arousal or intensity of

emotion. Keltner and Ekman (2000) also mention that the experience of different emotions results in varied autonomic responses. For example, sympathy is not only associated with oblique eyebrows and concerned gaze, but also heart rate deceleration, whilst distress is related to an increased heart rate, as is laughter. However, the latter is also linked to an elevated respiratory response. In addition, there are also varied responses or behaviour that can be elicited owing to the observation of varied emotions. For instance, facial expressions of anger can induce fear autonomic responses and facial expressions of distress may induce sympathy associated reactions.

Keltner and Ekman (2000) suggest that a dimensional view may be more appropriately employed for analysing emotion across time and moods. In addition, Keltner and Ekman (2000) discuss the reliability of using facial expressions as self-reports, which has been experimentally demonstrated to be consistently significant, but has shown statistically to be small to moderate (Matsumoto 1987). However, the use of facial coding systems has proven to enhance the relationship between the context of the experience and a person's self-report (Ekman et al. 1980). However, this accuracy can differ if the participants lie about their experienced emotional state.

According to Norman et al. (2003), affective and cognitive procedures happen in parallel in animals and people in order to process information. The latter comprise mechanisms that understand, interpret, remember and reflect upon information about the world. The former are mechanisms that assist in performing an immediate evaluation of events, i.e. assign a valence or value such as positive or negative, with respect to the person. These procedures are deeply interrelated. Affect is important because without affect there is no intelligent behaviour. Research has shown that people with brain damage in the area comprising the affective system become incompetent in complex world scenarios, e.g. setting goals and assigning priorities and are incapable of showing intelligent behaviour.

Norman et al. (2003) describe how cognitive and affective processes interact in order to produce behaviour through their three level model shown in Figure 2.2. The Reaction level is where the lowest-level processes, which are genetic, are executed in order to give immediate responses to information that comes from the sensory system, e.g. monitors the environment and the individual. These processes are capable of interrupting the execution of high level processes in order to give an immediate response to the stimulus.

The routine level is where all highly skilled and learned behaviours reside, e.g. language generation. These processes have access to short and long term memory and evaluation and planning methods. The information employed by the routine level comes from the sensory system and the reflection level. The reflection level comprises meta-processes that take decisions and generate control signals that are employed by the routine to decide if it has to activate or inhibit behaviour. The reaction level can pass information to the reflection level

when inconsistencies in norms and expectations are identified. The problem involved in using behaviour to identify the experience of an affective state is that one behaviour can be related to more than one affective state (Ortony et al. 1990).

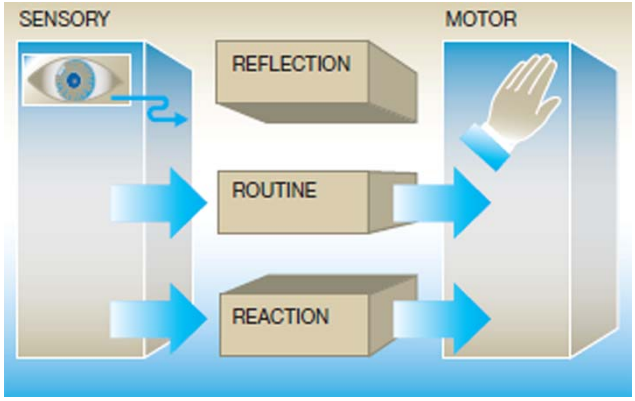


Figure 2.2 Three level model of human behaviour (Norman et al. 2003, p.39)

According to Brave and Nass (2008) there are three distinct forms of affect: emotion, mood and sentiment. Emotions are intentional and short term when compared with moods, their function is to prepare for an immediate response biasing action, contribute to moods and have a relationship with a specific entity, e.g. being angry with someone, while moods are longer lasting, bias cognitive processes, influence emotions and are unintentional, but may be caused by an entity. However, sometimes their cause is ambiguous and general, e.g. been depressed. Sentiments are assigned properties to an entity and are the result of previous experiences and generalisations and are the product of social interactions. They are not states. An example of sentiment is describing an interface as 'frustrating'. Ortony et al. (1990) mention that even emotions can be classified into positive or negative emotions, in order to distinguish systematically all qualitative differences between specific and individual emotions such as anger and relief.

This section reviewed different behavioural, cognitive, biological, neuropsychological viewpoints and the approaches that have been employed to gain insight into emotion with the main objectives of explaining it and categorising it. That emotion has physiological implications was also discussed. The latter topic is elaborated upon in the following section.

2.2 Physiological changes of emotion

Emotion generates physiological arousal. Physiological arousal can be monitored using biomedical electronic equipment (Schwartz and Olson 2003). The monitoring of physiological signals is related to the Biofeedback domain. There are two physiological characteristics that have been shown to be highly related to the experience of affective state: heart rate (HR) and galvanic skin response (GSR) or electrodermal activity (EDA). The operation of the

sympathetic nervous system is influenced by a change in arousal, i.e. reaction to stimuli. As a result, the body also experiences a series of physiological changes or processes: (1) muscle tension, (2) peripheral vasoconstriction and (3) altered electrodermal activity (Peek 2003, Rajae-Joordens 2008). These three characteristics, particularly the first two, are considered the most common biofeedback modalities, i.e. key channels where biofeedback signals are transmitted. These biofeedback modalities have also been shown to be strongly associated to anger, fear and excitement and electrodermal activity has also shown to be related to mental states, such as concentration.

According to Peek (2003), a biofeedback instrument, which is used to acquire and visualise biofeedback characteristics, has to serve three main functions: monitoring a relevant physiological process, provide an objective measure of it and present meaningful information about it. Peripheral vasoconstriction causes changes in heart rate, since dilated vessels pass more warm blood than constricted vessels do. These changes are perceived more easily in extremities, e.g. fingers and toes. As a result, peripheral temperature and phototransmission are indicators of this physiological characteristic. The latter is related to knowing that less blood in extremities, such as a finger, allows more light to pass.

On the other hand, GSR correlates with the activity of sweat glands, since sweat is comprised of conductive salts. As a result, sweaty skin conducts more electricity than dry skin. If conductivity increases in the skin is likely that this is due to the activation of more sweat glands. In order to measure GSR, the biofeedback instrument has to pass a voltage through the skin, e.g. the surface of fingers or place where many sweat glands are located, and it measures the amount of current allowed to pass by the skin. The measure can be a raw value or an objective measure of electrical conductance in μmhos. The advantage of objective measures is that they can be directly compared between individuals. Raw values are relative indications, since the scale is arbitrary and not standardised. It is observed that measuring peripheral vasoconstriction and sweat gland activity directly is not feasible. Therefore, biofeedback instruments measure accessible correlates that indicate the existence of a physiological process.

Rajae-Joordens (2008) employed GSR and HR to evaluate the experience of users while interacting with video games and watching movies using 2D and 3D TV screens. Results showed that whilst playing games, those interacting with 3D TV screens showed higher levels of GSR and higher scores of engagement and sense of presence than when using 2D screens. They did not find significant differences in the HR signal. Therefore, it was concluded that GSR is more effective for identifying engagement than HR. In addition, average HR and average GSR showed to be correlated to students' level of self efficacy as observed in the work by McQuiggan et al. (2008) and GSR has been employed in educational settings to identify students' affective state as can be observed in the work of Arroyo et al. (2009).

Figure 2.3 illustrates the electrical model of the skin illustrating its main features. The skin of the surface of the palm or fingers is comprised of approximately 2000 sweat glands. Each sweat gland is considered as a separate resistance or pathway. For convention, it is considered that current travels from a higher or positive potential to a lower or negative potential. The surface of the skin has higher resistance than deep layers of the skin. When a sweat gland is activated, it is said that it is in an 'on' state, creating a path from high-resistance to low-resistance, where current can flow. If the sweat gland is deactivated or in an 'off' state, this path is broken and current cannot flow on it.

As a result, if two electrodes are connected over the skin and voltage is applied through them, a closed circuit will be created, in which current can flow. The size of the current depends on the skin resistance or the number of glands turned on. The conductance is the reciprocal of the resistance. Conductance is employed instead of resistance, since it increases in a linear relationship with the number of activated sweat glands, while resistance decreases in a non linear relationship. Gasperi (2010) describes a method for building a GSR sensor using the LEGO intelligent brick and using a LEGO Mindstorm NXT program to acquire the raw value of the GSR sensor and display it as a numerical value (between 0 and 1023) on the NXT brick display.

Figure 2.4 illustrates the features of a GSR signal. The *tonic level (SCL)* corresponds to a baseline, i.e. resting level. It is a measure when a person is inactive and represents a baseline of gland activity or sympathetic arousal. *Phasic changes (SCR)* are episodes of increased conductance owing to arousal, which are easily distinguishable, e.g. it is observed an increase in conductance that peaks, levels out and return to the *tonic level*. Its amplitude is related to the degree of arousal caused by stimuli and it is calculated as the number of units above the *tonic level*. The *SCR half-recovery time* corresponds to the time elapsed from reaching the peak in a phasic change to one-half of the way back to the *tonic level*. It is an index of a persons' capability for calming down after an excitation. The *SCR latency* is the time elapsed from the application of a stimuli to the beginning of the rise time of a *phasic change (SCR)*.

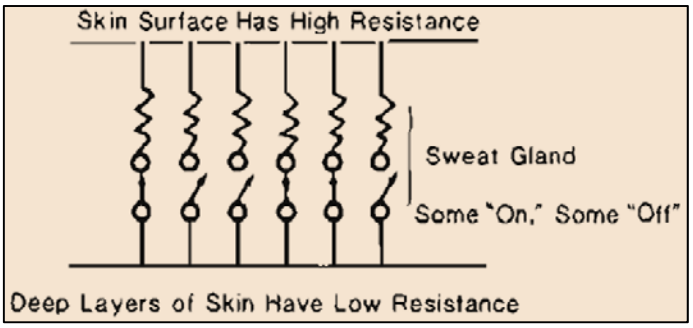


Figure 2.3 Electrical model of the skin (Peek 2003, p. 71)

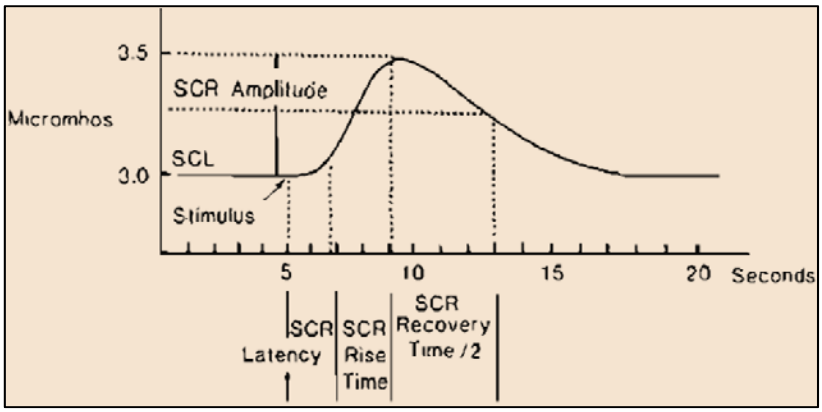


Figure 2.4 Features or parameters of a GSR signal (Peek 2003, p. 75)

Changes in sympathetic arousal influence HR, which is related to variations in peripheral vasoconstriction that impact body temperature, especially in extremities (Peek 2003). This is owing to the rate of blood supply. Dilated vessels pass more blood than constricted vessels. The measurement of peripheral vasoconstriction, i.e. vascular diameter, is inaccessible. However, the temperature of peripheral tissue surrounding the vessels can be measured using a biofeedback device. It can be measured in Farenheit (F°) or Celsius (C°) degrees. Also, variations in blood volume can be measured though photo-transmission, where more light passes through extremities with less blood and vice versa. However, it is important to remember that the amount of light that passes through the skin also depends on the coloration of the skin, e.g. pale or dark. The light intensity corresponds to a specific electrical signal. A *photoplethysmograph* is a device employed to measure the pulse. An average of the pulses can be obtained indicating the relative blood volume.

A way of creating a test of temperature is using *thermistors,* which are temperature probes, i.e. an electrical insulating material with wires employed for making thermal contact with an entity, e.g. the skin of a finger, for acquiring its temperature (Peek 2003). Usually the probe incurs to some extent in a delay before showing the actual value of temperature of the skin of the specific extremity (toe or finger). Figure 2.5 shows the theorised diagram of a temperature feedback device. The functioning of this heat-sensitive probe follows Ohm's law and can be understood using as analogy the operation of a water valve, where the voltage, the pressure in volts, pushes the current in amperes or water flow through the resistance in ohms of the circuit or pipes. The electrical resistance of the probe decreases as it heats causing more current to flow (the valve opens), whilst when the probe cools, the resistance increases provoking less current to flow (the valve closes). As a result, the amount of electrical current flowing corresponds to the value of the skin temperature.

This section discussed emotion that results in physiological changes, e.g. muscle tension or changes in peripheral vasoconstriction or electrodermal activity. Regarding the final two cases, two biofeedback instruments were presented and their functioning explained. It was

shown that since these changes cannot be directly measured, biofeedback devices measure their correlation. The following section examines emotion in an educational context.

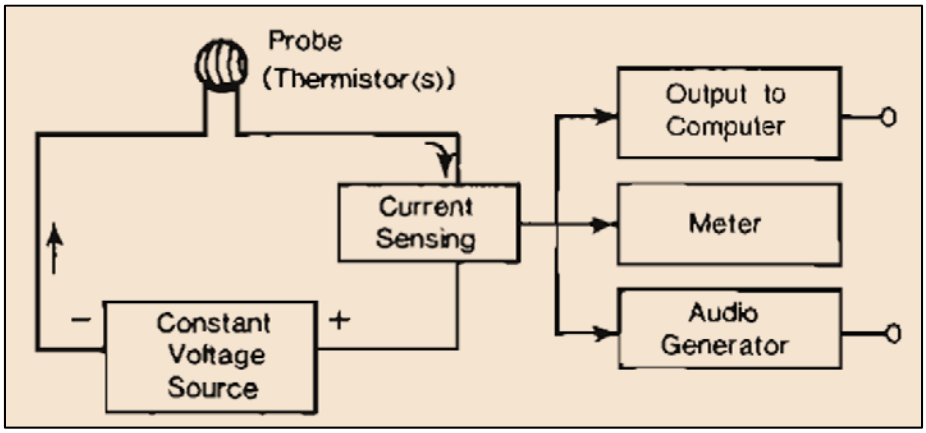


Figure 2.5 Temperature feedback device (Peek 2003, p. 64)

2.3 Emotion in education

In educational settings, emotions have been shown to influence student interactions, learning, understanding and performance (Pekrun et al. 2007). Also, emotion influences students' self-regulation and external regulation of learning, the strategies applied for achieving the learning objectives, students' motivation and the cognitive resources available. However, it is still not clear what the relevant emotions to the educational context are, which cognitive elements/factors participate in determining these emotions, or what the exact effects of emotions on students' learning experiences are. Educational approaches have been in constant evolution and are recently attempting to account for students' emotions. First, this section centres on presenting a concise explanation of the progression of these educational approaches, and finally focuses on current research that endeavours achieving an enhanced understanding of the causes and effects of emotions in educational settings.

Education has focused primarily on understanding how students' learn in order to enable learning to occur (Fry et al. 2009), since not all the students learn in the same manner. Whilst attempting to achieve this goal, several approaches have arisen. *Behaviourism* conceptualises the student as a passive being, who learns through positive and negative reinforcement and responds to stimuli (Learning Theories 2012). The final result is a change in student behaviour. Outstanding supporters of the Behaviourism approach are Pavlov (1927) and Skinner (1950). *Cognitivism* conceives the student as an active-rational being, i.e. a participant, and focuses on studying how students' mental processes happen, such as thinking, memorisation and problem solving. From the process of learning a change is expected in the stu-

dents' schemata and in turn a change on student behaviour. Key advocates of Cognitivism are Gagné et al. (1992) and Schank (1996, 1994).

Constructivism envisages students as independent and active learners who build new knowledge using previous knowledge, e.g. experiences, hypotheses and social factors. The final result is amended student schemata. Two of the key advocates of constructivism are Piaget (1950, 1973) and Bruner (1960). In particular, Bruner (1960) has a strong influence on changing the instructional approach from memorisation to understanding. Some of the techniques employed to assist students assimilating and transforming knowledge are reflection and experiential learning. The objective is to help students to retain knowledge for the long term.

Social theories of learning have also arisen, such as *situated learning*. The latter considers that student knowledge is the result of social practice (Lave and Wenger 1991). It is about enabling the student to achieve knowledge and understanding through practice in a specific context. *Experiential learning or learning-by-doing* is a type of instruction that it is achieved though action and work, which results in experience gained by an individual on his/her life (Fry et al. 2009). It considers that knowledge is in constant modification and renovation, i.e. a learning cycle, to arrange and include specific experiences. The main exponent of experiential learning is Kolb (1984), who contributed the *Kolb Learning Cycle* (Figure 2.6). Students involve and participate voluntarily in experiences (*concrete experiences*). Afterwards, students require a specific period of time to reflect upon these experiences from different viewpoints assisted by the feedback from others (*reflective observation*). From this reflection, students may achieve new ideas, which will be rearranged between pre-existent knowledge as theories or hypotheses (*abstract conceptualisation*). Students will use their understanding to take the decision that may solve the problem in the most appropriate manner. This decision will be tested (*active experimentation*), giving the opportunity of participating again in new experiences.

A theory related to Constructivism and Experiential learning is *Constructionism*, also known as 'learning-by-making'. This theory argues that learning happens as a result of the creation of meaningful objects in the physical world that involve the handling of materials. The two main proponents of *Constructionism* are Papert and Harel (1991). Two other teaching strategies related to *experiential learning* are Problem-based learning and Project-based Learning. In the former students are expected to understand and assess the nature of a real-world problem and, as a result, to find a suitable solution (Boud and Feletti 1997), i.e. the exploration of a scenario facilitates learning. The lecturer has a facilitator role in this situation, since his/her main duty is to provide material and facts that help students to gain insight about the problem. Students develop skills such as critical thinking and learn to work collaboratively. In Project-based learning, students follow a research driven approach, they seek

knowledge that can be applied to solve real-world problems (Dickens and Arlett 2009). Students develop their knowledge-based skills. Project-based learning is an approach well suited to applied topics, which offer multiple solutions to a specific problem. When using a Project-based learning approach, the lecturer's role is to assist students in becoming independent learners.

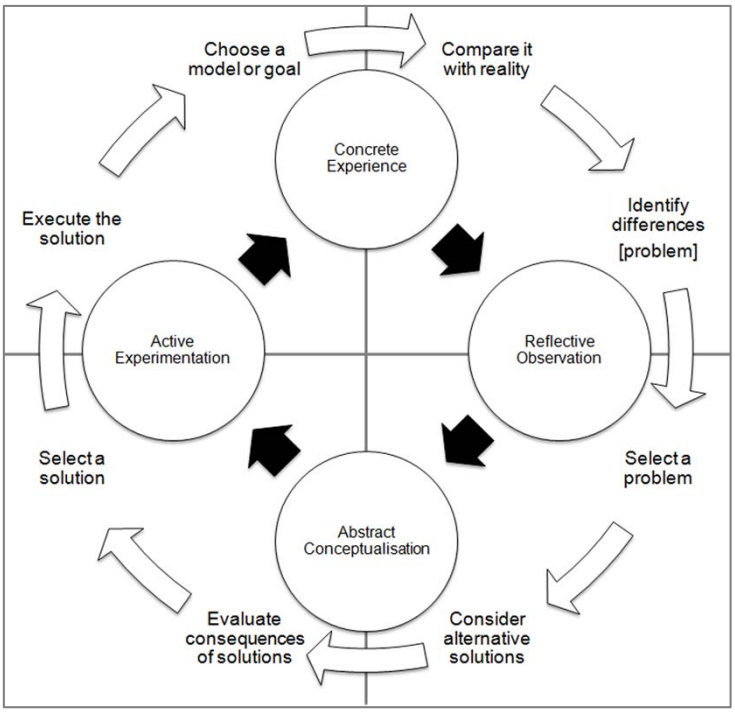


Figure 2.6 Kolb's learning cycle (Kolb 1984, p. 33)

Research has shown that students' final intention when approaching a learning activity determines his/her motivation, engagement and the quality of the achieved outcomes (Fry et al. 2009, Marton and Saljo 1984). This approach, situational and personal, may be a *deep approach*, a *surface approach* or a *strategic/achieving approach* to learning. When students employ a *deep approach* to learning, it is highly probable that the origin of the learning activity was intrinsic and self-imposed. Therefore, their goal is to achieve meaning, evaluate and identify relevant concepts and relations between new knowledge and previous knowledge. It requires high cognitive levels of processing and, as a result, high availability of time to invest in such activity. In a *surface approach* to learning, students' final goal is complete the task, which can be achieved through memorisation, even if this does not assist to derive new knowledge. It is very likely that the origin of this learning activity was extrinsic and imposed by others. A strategic/achieving approach is utilised when students' final goal is a high assessment mark. Therefore, the student will employ a combination of *deep* and *surface approaches*.

Humanism conceives the student as a person with values, motivation, interests, goals and intentions that can change and develop over the lifetime. Therefore, a student has to be studied as a whole (Huitt 2009, Learning Theories 2012). Humanist supporters think that constructing meaning is key to learning. According to Huitt (2009), humanistic education focuses on understanding the development of the student's regulatory and affective systems that have been neglected by previous educational approaches. From the Humanist approach arises the interest on student emotion and affect. The regulatory system operates as a filter that connects internal thoughts, feelings or knowledge with the external world through actions. The affective system examines and evaluates the information acquired from the regulatory and cognitive systems in order to increase or decrease its relevance in order to take action.

Humanism is grounded in principles such as assisting students to learn and develop skills that are relevant to them, teaching students how to learn, encouraging student self-evaluation, considering feelings as important as facts and ensuring that learning happens in a psychologically and emotionally secure as well as physically harmless environment. Therefore, the humanistic educational approach has five key objectives (Gage and Berliner 1998): (1) encouraging independence and self-instruction, (2) acquiring the capacity to take responsibility for what is learned, (3) developing creativity, (4) developing curiosity towards learning and (5) developing an interest in arts. These objectives are specially defined to support the development of the regulatory and affective systems and promote exploratory behaviour.

The advent of the Humanistic educational approach has encouraged and joined research efforts on emotion. One of the theories, which has currently arisen with the purpose of understanding the origin of emotions in educational settings is the Control-value theory of achievement emotions (Pekrun et al. 2007). Control-value theory is an integrative framework that is based on assumptions from several theories such as expectancy-value theories of emotion, theories of perceived control and models addressing the effects of emotion on learning and performance.

A high-level overview of the Control-value theory is shown in Figure 2.7. Pekrun et al. (2007) consider specific achievement activities and their future and past outcomes (Figure 2.7 (1)). Therefore, *Achievement emotions* can be defined as emotions directly related to achieve relevant activities and outcomes, which may be experienced by any participant in the educational context, e.g. teachers and principals. Achieving is a quality that is judged or evaluated according to established standards of excellence. For example, activities involved during students' learning are related to behaviours and outcomes that are judged against quality standards. Therefore, emotions in educational settings are considered *achievement emotions*. In addition, social emotions can arise in parallel and may overlap with the achievement of others, e.g. envy or empathy.

Control-value theory states that students experience *achievement emotions* when they feel in control or out of control of relevant achievement activities and outcomes from the students' perspective. Therefore, control and value appraisals are considered the key cognitive constructs involved in determining these emotions. Control is related to students' beliefs about their capabilities to perform an activity and achieve an outcome, but also it is related to their actual skills and strategies. Value is related to the value of the activity or the outcome per se, and results from evaluating the activity or outcome according to student goals, e.g. avoiding failure or achieving success. Achieving success is related to following a *deep approach* while avoiding failure is related to following a *surface approach*. Failure is perceived as a negative qualifier while success is perceived as a positive qualifier. The activity and the outcome have to be relevant from the student's viewpoint. They have to be interested in achieving them in order to experience an emotion. Pekrun et al. (2007) mention that if one of the appraisals, control or value, is missing the emotion would not be generated.

Control and value appraisals may be influenced by antecedents, such as a person's achievement goals related to control and value beliefs (Figure 2.7 (2)); genetic and temperament factors (Figure 2.7 (3)); and educational environment factors, social environment factors and socio-historical context factors (Figure 2.7 (4)). Control-value theory also considers the effects of *achievement emotions* on student cognitive resources, motivation, strategies and external/internal regulation (Figure 2.7 (5)). These processes mediate student achievement (Figure 2.7 (6)). Also, Pekrun et al. (2007) consider in the Control-value theory that students' outcomes revisit their emotions (Figure 2.7 (7)) and the environment (Figure 2.7 (8)) closing the feedback cycle. Antecedents and effects of emotion are related by mutual causation over time (Figure 2.7 (1-8)). This mutuality has repercussions for the regulation of achievement emotions (Figure 2.7 (9-11)) and emotional learning environments (Figure 2.7 (12)).

Affect, i.e. emotions and moods, can be conceived from a 'trait-like' or a 'state-like' perspective (Linnenbrink 2007). The former believes that affect, which is a manner of responding to world stimulus, varies from one person to another. However, it is relatively stable over time. The latter also considers that affect is a response to the changing world. However, it believes that affect is strongly determined by the specific situation or context in comparison to the influence of personal differences. Affective states can also be differentiated through intensity and duration. Categories of affect can be defined based on dimensions (Linnenbrink 2007), such as *value* and *activation*. The valence is induced by events, which can be pleasing or unpleasing and the activation is related to the preparation for action, mobilisation, energy and arousal.

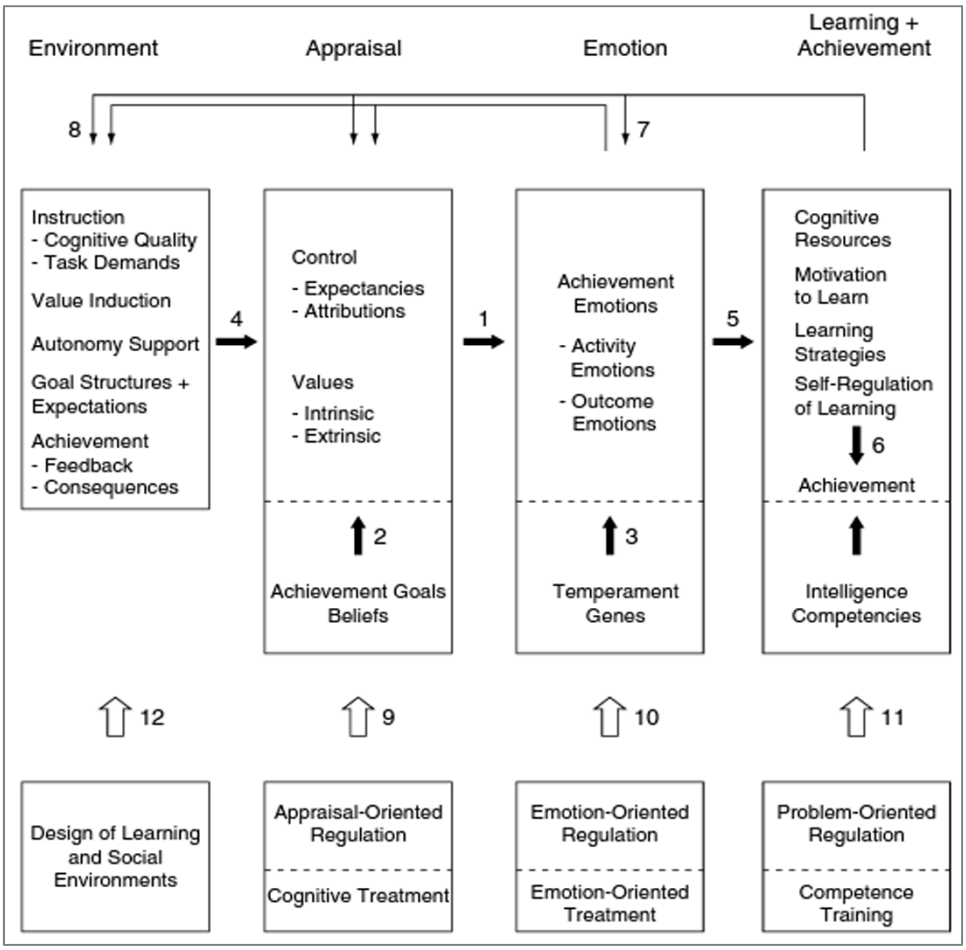


Figure 2.7 Control-value theory diagram (Pekrun et al. 2007, p. 17)

Feldman Barret and Russell (1998) crossed these two dimensions, value and activation, in order to define key affective states, such as tense, depressed, relaxed and excited, creating the *circumplex model of affect* (see Figure 2.8). This model also captures affective intensity. From the *circumplex model of affect* it can be observed that an affective state located in the Activation/Unpleasant quadrant has different implications for education than an affective state located in the Deactivation/Unpleasant quadrant (Linnenbrink 2007). Since, the former can lead to more student engagement than the latter, e.g. nervous against bored. In comparison with Control-value theory, the *circumplex model of affect* also considers moods and differences affective states using *value* and *activation* dimensions. Whilst Control-value theory distinguishes an emotion using appraisals and underlines the *object focus* and *valence* dimensions. *Therefore, achievement emotions* can also be classified according to these two dimensions: valence (e.g. positive or negative) or degree of activation (e.g. activation or deactivation), but also achievement emotions are classified using the *object focus* dimension, see Table 2.1.

Control-value theory proposes three kinds of achievement emotions: (1) prospective outcome, (2) activity and (3) retrospective outcome emotions. These are defined according to

the object in focus, e.g. outcome/activity, and the time frame of the emotion, e.g. before achieving an outcome, while performing an activity or after achieving an outcome. Therefore, undergoing activities and past and future outcomes are considered when experiencing and determining an emotion. The classification of achievement emotions and its relation with control and value appraisals is shown in Table 2.2.

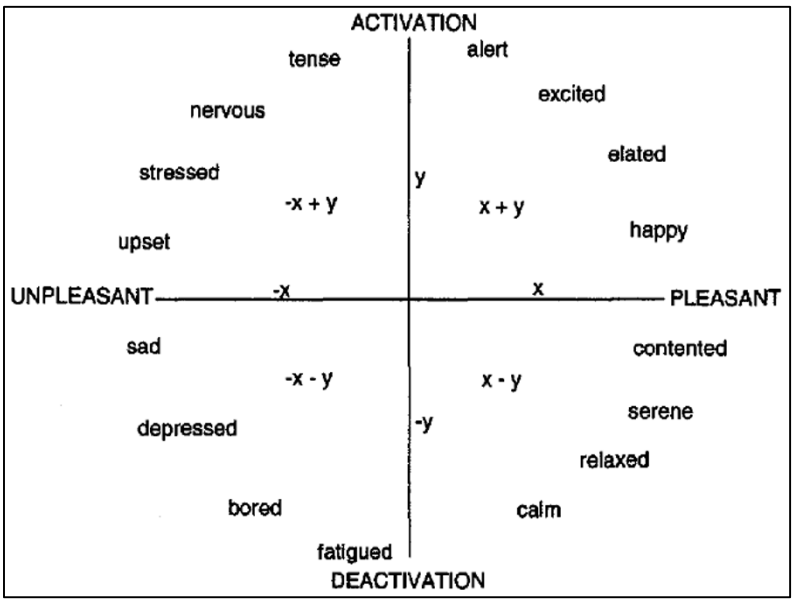


Figure 2.8 Circumplex model of Affect (Feldman Barret and Russell 1998, p. 970)

	Positive[a]		Negative[b]	
Object Focus	Activating	Deactivating	Activating	Deactivating
Activity Focus	Enjoyment	Relaxation	Anger Frustration	Boredom
Outcome Focus	Joy Hope Pride Gratitude	Contentment Relief	Anxiety Shame Anger	Sadness Disappointment Hopelessness

[a]Positive, pleasant emotion; [b]Negative, unpleasant emotion.

Table 2.1 Value and activation of *achievement emotions* (Pekrun et al. 2007, p. 16)

Prospective outcome emotions are experienced when failure or success are expected and also depend on the control that the students believe they have, e.g. high, medium or low. *Retrospective outcome emotions* are experienced when relevant success or failure outcomes were achieved and the control exercised over them is attributed to the self, to other agents or to the situation. Also there are emotions that are experienced at this stage that do not depend on the control exercised or high cognitive processes, such as joy and sadness. *Activity emotions* are related to the achievement of an activity and depend on the perceived controllability of the situation from the students' viewpoint, e.g. controllable corresponds to a posi-

tive evaluation. If the activity is not classified as controllable or uncontrollable, there is not an evaluation of the activity, e.g. none, boredom is experienced.

Achievement emotions are domain-dependent (Goetz et al. 2007). As a result, emotions experienced by students learning History are different from the ones that students can experience when learning Mathematics. In addition, Goetz et al. (2007) have also shown that emotions experienced in similar subject domains, such as mathematics and physics, have stronger correlations. Furthermore, each learning and achievement situation has a specific social structure and function. Pekrun et al. (2007) distinguish three kinds of situations: *test-related, class-related and learning related*, which also influence the kinds of emotions that can be experienced.

	Appraisals		
Object Focus	Value	Control	Emotion
Outcome / Prospective	Positive (Success)	High	Anticipatory joy
		Medium	Hope
		Low	Hopelessness
	Negative (Failure)	High	Anticipatory relief
		Medium	Anxiety
		Low	Hopelessness
Outcome / Retrospective	Positive (Success)	Irrelevant	Joy
		Self	Pride
		Other	Gratitude
	Negative (Failure)	Irrelevant	Sadness
		Self	Shame
		Other	Anger
Activity	Positive	High	Enjoyment
	Negative	High	Anger
	Positive/Negative	Low	Frustration
	None	High/Low	Boredom

Table 2.2 Classification of Achievement Emotions (Pekrun et al. 2007, p. 20)

Pekrun et al. (2005) created the *Achievement Emotions Questionnaire* (AEQ), a self-report tool to identify students' emotions in three different achievement situations. The AEQ comprises affective, motivational, cognitive and physiological factors and was designed through Structural Equation Modelling (SEM). The AEQ has been employed effectively to identify the emotions of undergraduate students of physics (Goetz et al. 2007). However, this theory has not been employed previously to derive or create a computational and emotional student model.

2.4 Summary

In this chapter, work in the area of emotions and education covering work in Cognitive Psychology, Biology, Behaviourism and Physiology was discussed. To reason about emotion, it is meaningful to understand what emotion is, the origin of emotions and their characteristics. Emotion is dynamic, short lasting and intentional. The term affect is employed to refer to emotions in addition to moods. Affect can be visualised as *states* caused as a result of changes in the situation, environment or context, or as a trait determined by a person's personality. Self-reports are considered evidence of emotions, since they are subjective and only the person has access to them.

Emotion influences students' motivation, strategies, learning and performance and plays a key role in students' physiology process and behaviour, the employed learning strategies and the use of the available cognitive resources. GSR and HR are two physiological signals that are highly related to individual's affective states. Also, it was observed that research has shown that GSR is more sensitive to, and is significantly associated with, changes in emotion.

Furthermore, to create a computational and emotional student model it is important to identify the relevant emotions to the educational experiences and how these are determined. *Achievement emotions* are emotions that occur in educational settings when students want to achieve relevant activities and outcomes, e.g. boredom and frustration. The *Control-value* theory of achievement emotions by Pekrun et al. (2007) assumes that control and value appraisals are the most relevant when determining these emotions and that these emotions are highly related to an achievement activity and its future and past outcomes. Also Pekrun et al. (2007) argue that there is a relation of mutual causation between the antecedents and effects of *achievement emotions* over time.

Chapter 3: Virtual and Game-based Learning Environments

Two types of online learning environments are examined here, virtual learning environments (VLEs) and game-based learning environments (GBLs). These environments enhance knowledge acquisition through offering flexible and interactive learning. Our discussion centres on the advantages that Intelligent Tutoring Systems (ITSs) provide in these environments followed by a review of student modelling.

Affective Computing is currently influencing research in two computer-instruction areas, namely ITSs and Edutainment. We investigate current ITSs, since they introduce and exemplify approaches and techniques employed to recognise or reason about students' emotion and personal disposition, e.g. self-efficacy, attitudes and motivation. Work on the recognition of personal disposition by ITSs is included due to the special relationship that motivation and cognition have with respect to emotion.

3.1 Online learning: Virtual & game-based learning environments

Students' expectations have changed, e.g. they are not easily engaged with traditional instruction methods and have a preference for learning through inductive thinking, which entails various flows of information and forms of interacting with content. This current student generation is known as the *Net generation* or *Digital Natives* (Oblinger 2004, Van Eck 2006). Therefore, online learning is a more plausible solution to meet current students' demands. Online learning is defined as access to flexible and interactive learning experiences through the use of technology (Moore et al. 2011). Open-ended learning environments (OLEs) support online learning through enabling students to experiment, interpret and learn from their mistakes. VLEs and GBL environments are examples of OLEs.

VLEs enable students' interaction locally or remotely through simulations or learning objects in a suitable combination of activities and delivery of content (Joint Information Systems Committee 2009, Reilly 2008). VLEs' main challenge, as in traditional instruction, is to determine the combination of resources and instructional approaches that constitute the most suitable delivery method. In addition, students easily grasp the link between real world concepts and phenomena with objects and events of the VLE. It is important to highlight that the error cost of learning to operate equipment or to perform a medical procedure in a simulation is significantly reduced and learners develop a feeling of security (Connolly and Stansfield 2006).

According to the form of instruction that it is required, the learning objects of VLEs can be designed as *self-directed*, *self-paced* or *instructor-led* (Moore et al. 2011). *Self-directed* refers to learning where students are in charge of managing and monitoring cognitive and contextual aspects of their own education. *Self-paced* refers to enabling autonomy for learning in the students' own time, peace and location. An example of a *self-paced* learning system is Khan Academy (2012), which includes a video library, interactive challenges and assessment activities. Another example is Udacity (2012), which also uses videos to assist students in learning concepts using real-world problems. The lectures are created by prominent instructors at University-level. Udacity provides students with the opportunity of receiving recognition for the skills attained.

Instructor-led refers to a form of learning where the instructor controls the learning-sequence, which involves the participation of learners in the same activities at the same time. Efforts here are working towards a more enhanced, autonomous and personalised learning experience, where satisfying identifiable students' needs is the main objective (Joint Information Systems Committee 2009). It is here that ITSs have played a key role in the achievement of this goal.

GBL environments also enable learning through experiencing the effects of students' own actions in situated contexts and facilitate the connection between learning and real-world experiences (Van Eck 2006). However, GBL environments have proven more effective in maintaining students' engagement than VLEs (Connolly and Stansfield 2006, Muñoz et al. 2009a). GBL works, since the act of playing is considered a primary instructional strategy and a form of socialisation. As a result, Edutainment, which is a sub-class of serious games, i.e. games employed for a serious purpose, easily engages students' attention and is focused on enabling students to play in order to learn and enjoy the experience of learning (Ma et al. 2011, Qianping et al. 2007).

In addition, GBL environments or educational games provide immediate feedback and reward learning and mastering through different modalities, e.g. heroic music, new characters, progression of story, a high score and power-ups (Sykes 2006). However, common problems encountered in GBL design are balancing game and learning content, supporting the curriculum and ensuring that learning actually happened (Carpenter and Windsor 2006, Conati 2002, Sykes 2006, Van Eck 2006). Edutainment also includes personalisation (Paireekreng et al. 2009), and hence ITSs are also used in combination with GBL environments in order to achieve personalised instruction and ensure the achievement of learning goals (Conati 2002).

3.1.1 Emotional GBL environments

The most important characteristic of GBL environments is arguably their emotional character, which facilitates the establishment of an emotional relationship with the learner (Sykes 2006). GBL environments are comprised of identifiable elements, e.g. narrative, characters, sounds, actions, challenges and goals. These elements interact creating a unique experience, known as gameplay (Rollings and Adams 2003). It is for this experience, emotionally loaded, that people play games (Lazzaro 2004). In addition, GBL environments comprise elements that are emotional in nature, such as narrative, sounds or music and graphics or animations.

In video games, narrative or storytelling sets a mood or communicates goals (Collins 2008). Additionally, according to Schank (1996) narrative serves a cognitive function, since it assists us in organising our personal experiences and achieving understanding of ourselves and the world around us, which is a statement that supports GBL environments instructional capability. However, it also can be argued that stories in GBL environments serve purposes of emotional intelligence, since they enable us to evaluate easily events or facts involved in complex issues. It is important to highlight that evaluation is a human skill owing to affective and emotional processes that operate in conjunction with cognitive processes, as is argued by Norman et al. (2003). In addition, stories enclose and communicate emotions, since usually they are the result of life experiences as can be observed in the following example (Schank 1996):

> "I've been busy selling my house and otherwise preparing for the move. All of a sudden, my appointment has been stopped at the highest levels of the company. No one will tell me why, but I think someone who was my enemy in the past has a friend at the company. And I think she wrote a letter that prejudiced them against me. I'm very upset" (Schank 1996, p. 29).

Music and sound in video and educational games are employed with several purposes. According to Collins (2008), sound in games provides feedback in response to players' actions and game states. Such sounds are known as *adaptive audio*. From an audio perspective, players are transmitters and receivers simultaneously. Also, audio in games enhances immersion, e.g. sounds related to the game atmosphere or to the game dialogue. Game genre influences the functions that audio must serve. Genre is an important feature that has been employed by the games and software industries to define categories of games with characteristics that are considered attractive by specific groups of users (Bateman and Boon 2006). Types of interactivity or interface with the player, narrative, representational and gameplay rules are characteristics of game genre that influence *adaptive audio* and *immer-*

sive audio. In addition, sounds are used to communicate emotional meaning or cues concerned with decision making, and to manage players' attention in order to distinguish goals, objectives or new characters (Collins 2008). Also, sound assists in creating a sense of continuity, which involves matching the emotion and pace of the game, and can be employed to induce a mood. As a result, game genre and sounds assist gamers in easily comprehending the gameplay, hence decreasing the time invested in learning to play successfully.

On the other hand, non-player characters (NPCs) are usually a combination of graphics or animations and incorporate AI techniques, which are usually developed to serve three main purposes: (1) signalling game goals, (2) taking part in the storytelling process and (3) engaging us in the game story (Aqel et al. 2009). Toward these goals, these agents or synthetic characters may show attitudes, moods and emotions, which are part of character personality. Additionally, according to Freeman (2004) perhaps the key to influencing player emotions is to create characters with whom players identify and which express authentic moving experiences. However, these two endeavours still represent challenges, since it cannot be assumed that players will experience an emotion, as games follow a non-linear progression, i.e. events and timing cannot be completely controlled, and players may not always feel affinity towards characters, since it depends on player background, e.g. culture, social standards, goals and attitudes.

Colours are also a feature of NPCs and the surrounding environment, according to Block (2008), is a *weak visual component*. However, they can be employed in terms of affinity and contrast to convey meaning, e.g. express anger, show corruption or indicate danger. As observed, colours may be related to one or various meanings (Zammito 2005): (1) innate, (2) personal and (3) cultural. Innate meanings are associated with our evolution process, e.g. surviving and adapting signals. Personal meanings are linked to our own experiences and cultural meanings are acquired from our interactions with society, which implies sharing to some degree the same environment, feelings, moods and thoughts. As a result, colours are capable of influencing our emotions and mood (Kaya et al. 2004) and sometimes they encourage us to act (Zammito 2005). Colours give us the sensation of contact or removal from our environment. For example in Figure 3.1, a car approaching in the street in a warm colour, such as red or orange, may be perceived as being nearer than the same car in a cold colour, such as blue or purple, which will be perceived as being farther away. Also, the tone of colours may be employed to handle players' attention, since the brightest areas usually are more attractive (Block 2008). Furthermore, saturated colours are frequently associated with positive emotions, such as happiness, while desaturated colours are usually associated with negative emotions, such as sadness.

Figure 3.1 Contrast between warm and cold colours

3.1.2 Overcoming design problems in GBL environments

Learning to play an educational game successfully does not necessarily ensure that students have grasped the required knowledge (Conati 2002), which is related to incorporation of measurable learning outcomes and an intelligent mechanism of assessment, such as an ITS, in the implementation of a GBL environment. In addition, achieving an appropriate equilibrium between game content and domain knowledge may prove challenging (Carpenter and Windsor 2006, Sykes 2006). Design problems can bias educational institution or lecturer opinions against incorporating GBL environments in their curriculum and not investing in their acquisition or development. Therefore, research has focused on defining design principles (Squire 2006, The Games-to-Teach Research Team 2003), analysis and design methodologies and models (Akilli and Cagiltay 2006, Bateman and Boon 2006), and understanding which elements of gameplay encourage learning and engagement (Collins 2008, Malone and Lepper 1987). Creating and designing more successful GBL environments has entailed analysing previous work in the video game industry (Games-to-Teach Team 2003) and analysing learner interaction patterns (Akilli and Cagiltay 2006). In a similar manner, work focused on understanding the effects of technology over targeted populations (Chumbley and Griffiths 2006).

According to the Games-to-Teach (2003) research team – constituted by members from Microsoft and the Massachusetts Institute of Technology (MIT) – narrative in educational games should be compelling and designed in order to set the context of the activity, introduce goals and constraints, encourage thinking and shape actions. Design principles suggested by Games-to-Teach (2003) that, based on our experience, are relevant to achieving effective GBL environments are:

1. *Begin the creation of GBL environments by using simulations as a basis*, since even simulators model realistic phenomena and environments and GBL environments have also to incorporate elements of fantasy, such as an additional story and context. However, both need to represent specific inputs and outputs, include

one or more goals and may have various end states. Figure 3.2 summarises these similarities and differences in the design of VLEs and GBL environments.

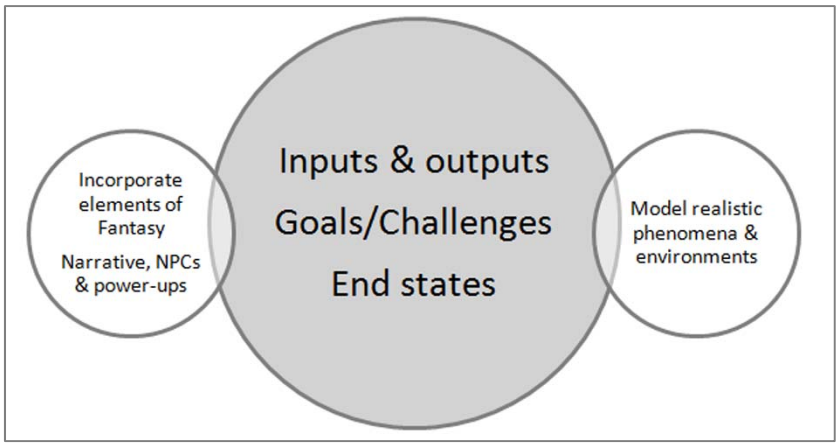


Figure 3.2 Similarities and Differences between VLEs and GBL environments

2. *Include in GBL environments educational content that encourages thoughtfulness in order to achieve meaningful and compelling goals*. Exemplifying this principle, adventure and role-playing games are naturally structured in a problem-based, goal-based or case-based scenario basis, where the achievement of challenges drives the progression of narrative. These challenges should encourage students to use their skills and knowledge in order to overcome them.
3. *Facilitate interesting choices and consequences*. This principle is related to designing various paths to overcome challenges and constraints by creating game rules that are linked to potential students' actions and decisions employed to solve those challenges. Sometimes students' actions may be linked to students' misconceptions. However, it is important to allow students to experience the consequences of their own errors in order to achieve learning (Squire 2006).

In addition to the derivation of design principles for creating effective edutainment or GBL environments, work on the conception of Instructional Design and Development Models (IDDMs) offering a paradigm that assists HCI to make the teaching process more efficient. The process is comprised of several stages, such as analysis, development and evaluation. An IDDM for creating GBL environments is the Fuzzified Instructional Design Development of Game-like Environments (FIDGE) model (Akilli and Cagiltay 2006). The FIDGE model was derived from real data using Fuzzy Logic, and is therefore realistic and does not follow a linear flow of execution (Akilli and Cagiltay 2006). The FIDGE model is comprised of pre-analysis, analysis, design, development and evaluation phases, which have vague boundaries, may be conducted in parallel or are occasionally revisited. Additionally, the FIDGE model can be used by novices or experts, since the pre-analysis stage provides a reference

for novices, e.g. assisting them in gaining insight on how to achieve awareness of learners' and lecturers' needs. Also, since students' expectations for learning are changing, this model assists lecturers in being aware of these changes and of the latest resources, technology and tendencies.

Bloom's taxonomy, which is a model that explains how we learn, is employed to achieve an effective design of pedagogical material and activities using learning goals (Bloom et al. 1984). Its main objective is to assist students in applying knowledge in order to solve real world problems. Bloom's taxonomy is organised into six levels: knowledge, comprehension, application, analysis, synthesis and evaluation. Recently, Bloom's taxonomy has been employed to solve the problems of game-design (Honeycutt 2011). It states that a design should start by defining the learning outcome and objectives. However, it is important to remember that in addition to game outcomes, a game's design can be based both on the process of playing the game and the process of debriefing/reflection. In the knowledge and comprehension levels of Bloom's taxonomy, the instructor is seen as a leader with tight control over the design process. The instructor chooses and directs the game-play: activities, learning content, feedback, help, guidance and support provided. In the application and analysis levels of Bloom's taxonomy, the instructor has a facilitator role, i.e. he/she has loose control over the design process. The instructor designs the learming/game content, but does not provide the game solutions if students do not explicitly ask for them and enables students to interact at their own pace. Students assess their own level of understanding. In the synthesis and evaluation levels of Bloom's taxonomy, the instructor has an observer role where he/she does not provide support in any way. He/she provides rules and an overview of the game. Players work independently or collaboratively to work out frustrations and disagreements. Players also lead the discussion about problem solving strategies. Instructors facilitate a debriefing session that depends on student input.

As mentioned earlier, game genre is an important feature from a marketing and educational viewpoint. Bateman and Boon (2006) suggest that according to personal preferences, people are more predisposed to choose and enjoy playing games of a specific genre or type of game play, which is important to take into account when designing games. They propose the Demographic Game Design model (DGD1), which entails four play styles – *wanderer, conqueror, manager and participant* - that are related to player expertise and preferences, e.g. casual or hardcore, and are also associated with Myers & Briggs personality types. Bateman and Boon's (2006) findings may assist commercial video game companies, universities and lecturers to understand the kind of games that are considered interesting from the target audience's viewpoint.

Educational and commercial games aim to achieve immersion, which entails users feeling part of an imaginary world and achieving emotional attachment (Collins 2008). According to

Glassner (2004) players can experience different degrees of immersion: (1) curiosity, (2) sympathy, (3) empathy and (4) transportation (see Figure 3.3). These states are arranged from left to right from the lowest level of immersion or engagement to the highest. Glassner (2004) explains that if the player or learner feels a necessity to know, it is said that he or she is in a state of curiosity, whilst if the player or learner begins to adopt the protagonist's viewpoint and judges the actions of NPCs and events in the virtual world according to this character's perspective, it is said that the learner or player is in state of sympathy. Otherwise, if the learner or player develops a feeling of identification with the protagonist through discovering that personal features are shared between him or her and the character, it is assumed that the learner or player is in a state of empathy. Finally, the learner or player can achieve a state of transportation when from the player or learner viewpoint the existing boundaries between himself or herself and the protagonist disappear.

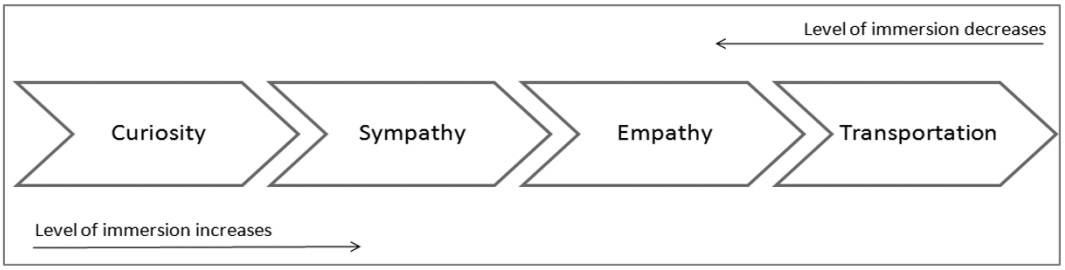


Figure 3.3 Levels of immersion (Glassner 2004)

Game immersion may be achieved in different manners (Collins 2008): (1) overwhelming real world information, e.g. using large screens very near to player or learner faces with high intensity sounds, (2) achieving right flow or balance between challenges and the development of skills to overcome them, (3) encouraging imagination through storytelling and NPCs, e.g. enabling players or learners to use their imagination, to empathise with and enjoy the fantasy and (4) removing users' disbelief through the construction of believable and realistic virtual spaces and NPCs. A major constraint on achieving user immersion is that it depends on the learner or player state of mind. As a result, it is not permanent and fluctuates whilst playing.

Different methods for influencing user immersion agree with the derived theory by Malone (1981) for making instruction intrinsically appealing. Malone (1981) signalled that challenge, fantasy and curiosity, which are characteristics of video games can encourage learning in these environments. Learners need clear goals that must be uncertain and relevant not only from an educational perspective, but also from the game fantasy viewpoint. It is important to remember that not all fantasies appeal to all users since personal preferences and gender influence user choice. However, fantasy or storytelling is important because users have con-

tact with other contexts that assist them in achieving an enhanced understanding of the situation.

Curiosity is highlighted by Malone as “the most obvious intrinsic motive for learning” (Malone 1981, p. 356) and it entails providing an optimal level of information, e.g. not giving unnecessary knowledge. The learner should not have the whole comprehension of the facts, otherwise novelty will be lost. Also, curiosity can be created through finding ways of attracting human senses (Malone 1981), e.g. a bright light or colour, a loud sound or an animation. According to Malone and Lepper (1987) control is another feature that influences students’ intrinsic motivation in order to learn and/or keep playing. It is in our human nature to endeavour to achieve a state of competence and seek self-determination, i.e. achieving control of our own environment through exercising actions and experiencing consequences. With regard to the context of VLEs, control is influenced by the available outcomes and the probability of achieving such outcomes. The learners’ insight of control and competence can be influenced by achieving a successful or failed outcome, since outcomes are the result of student actions and can prove to be beyond student capabilities. Therefore Malone and Lepper (1987) suggest adapting the level of difficulty of the problems presented, the feedback provided and the style of teaching employed according to student skills. Also, they suggest creating environments where students can choose the problem to be solved or its level of difficulty.

Malone and Lepper (1987) also signal interpersonal motivators for learning - competition, cooperation and recognition. Challenges have goals, a quantifiable manner of knowing if outcomes are fulfilled. Cooperation and competition can be explained in terms of utilities that are assigned to outcomes, which are the result of student actions. In competition utilities can be added giving a result of zero, e.g. loser or winner, while in cooperation if utilities are added they do not give a result of zero. To create environments that encourage cooperation and competition, the tasks should be designed as a series of activities conducted independently or dependently towards achieving the same goal. Recognition can be described as the enjoyment experienced through being acknowledged by others for the effort invested in our achievements (Malone and Lepper 1987). Recognition may be attained in the implementation of GBL environments through creating activities that are in some way observed by others whilst students are performing them or publishing activity outcomes.

GBL environments or educational games have to pursue personalisation in order to fulfil student skills and needs, otherwise learning cannot be more successful (Virvou et al. 2006). Personalisation entails adapting interaction to student actions in order to enhance student competence. To support this, personalisation is a necessary feature (Virvou et al. 2006, Woolf 2009), which should appear in GBL or edutainment environments. There are various approaches to personalisation in GBL environments (Bakkes et al. 2012, Woolf 2009), including space adaptation, mission/task adaptation, character adaptation, game mechanics

adaptation, narrative adaptation, music and sound adaptation, player matching, difficulty scaling techniques and ITSs. The next section focuses on discussing the main characteristics of the ITS approach for computer tutoring - especially in the area of student modelling - and finally focuses on its ultimate endeavour, which is related to reasoning about student emotions and addressing them appropriately.

3.2 Adaptable and personalised computer tutoring

ITSs are a form of computer tutoring that is based on a student-centred approach, where AI techniques are applied to modelling and reasoning about students' characteristics, skills, behaviour or needs over time and respond accordingly to them (Woolf 2009). These models can be derived with concepts that students should know and understand, misconceptions, learning preferences and psychological and cognitive theories that explain how students acquire knowledge - *student modelling* - and how lecturers diagnose learning - *tutoring expert modelling*. Advantages of ITSs enable students to advance at their own speed, achieving and engaging in their own learning. Also, ITSs are employed, since they have assisted in gaining more insight about the processes involved during instructional interrelations and facilitating one-to-one tutoring that has proven as effective as, and less time consuming than (Nwana 1990, Woolf 2009), one-to-one human tutoring. Learning experiences in ITSs can be designed to be self or group directed.

According to Woolf (2009) and Du Boulay and Luckin (2001) motivators of ITSs are: (1) changes in the teaching and educational methods, which are currently more focused on understanding cognitive processes, learning styles and interaction methods through observation of teaching-learning experiences, (2) changes in IT, e.g. Internet, which have encouraged novel ideas for processing and saving information, software development and the creation of networks, (3) advances in AI techniques, since these are employed to provide adaptable instruction through reasoning about student behaviour and interaction, managing appropriately resources, evaluating student learning or encouraging collaboration, (4) high availability of student data to be mined, i.e. Educational Data Mining (EDM), which can be employed to achieve an enhanced understanding of students, (5) progress in HCI paradigms related to GUI design that facilitate achieving learning and interaction goals, support new interaction techniques and keep track of and record data related to student interaction.

ITSs differentiate from Computer-Assisted Instruction (CAI) or framed-oriented systems through key features (Regian et al. 1996): the capability of *generating* appropriate challenges and adapting feedback to student needs and skills; the ability of representing and reasoning about student characteristics, knowledge, needs and interaction patterns and responding accordingly; the capacity of representing and reasoning about expert performance in the domain; the capacity of initiating, interpreting and responding to interactions with students;

comprising learning activities that encourage student participation and are properly contextualised and domain-relevant; the capacity of adapting teaching strategies and the ability of self-improving instructional performance. From these features, it is noted that types of knowledge for underpinning ITSs are *domain*, *student*, *tutoring* and *communication* knowledge (Woolf 2009) and that ITSs comprise research in three major areas, *Computer Science* (namely AI), *Psychology*, particularly cognition and *Education and Training* (Nwana 1990).

From the features previously mentioned arise the problem of how the architecture of ITSs should be constructed. There is a diversity of architectures, but there are four key components in all of them (Nwana 1990): GUI, expert knowledge, student model and tutoring modules, which are shown in Figure 3.4.

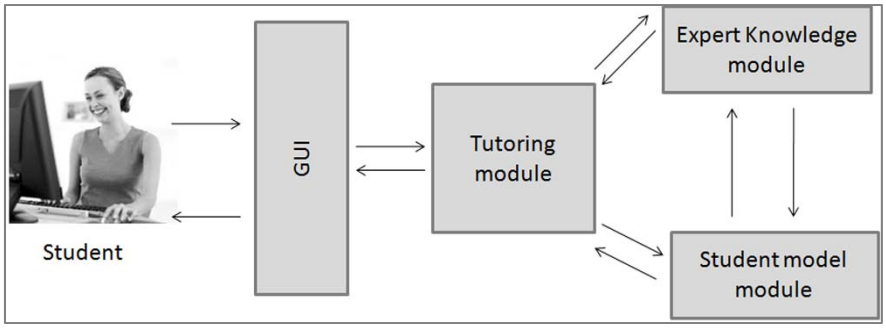


Figure 3.4 Generic architecture of an ITS

The GUI handles communication between the student and the ITS, which entails translating the ITS's internal representation to language that can be comprehended by students and vice versa. The expert knowledge module includes domain rules and domain facts that are selected by experienced people and will be taught to students. This knowledge is coded and represented to be exposed in specific situations where students require guidance. This process entails considerable processing time and depends also on domain complexity.

The student model module comprises a representation of student knowledge and skills under development and also includes student behaviour that is closely related to consequences in their performance and learning. The student model is also strongly linked to the available channels of communication between the student and the personal computer (PC). The tutoring module or teaching strategic or pedagogic module controls and manages educational interactions through using the available knowledge about the students and teaching goals to select the challenge or domain knowledge that will be presented.

ITS research focuses mainly on solving two problems: (1) how to achieve an enhanced reasoning and understanding of student needs, behaviour and skills and (2) how to achieve student learning, understanding and more recently student engagement and motivation. As a result, investigation has been principally conducted on the GUI and knowledge representa-

tion, e.g. *student, domain, tutoring and communication*. The latter involves applying AI techniques to modelling educational interactions using as a basis the observation of real one-to-one teacher-student interactions or cognitive psychological theories and methodologies (Alexander and Sarrafzadeh 2008, Du Boulay and Luckin 2001), which have attempted to explain the meta-processes involved in students' reasoning and acquisition of knowledge or the framework in order to achieve them, e.g. Socratic Style. Our investigation centres on the first challenge, i.e. *student modelling*. An ITS is a type of user-adaptive software system, which has the main characteristic of adapting interaction to each user and his/her needs (Brusilovsky et al. 2007). To achieve adaptability, a typical ITS will employ a generic representations of the user, which is known as *user model*. Hence, *user modelling* is the method through which user models are created and sustained.

3.2.1 Student modelling

As observed in the work of Alexander and Sarrafzadeh (2008), Kort et al. (2001) and Lepper et al. (1993), in one-to-one teaching-learning experiences lecturers require continuous observations to effectively identify student characteristics and preferences related to student performance and learning - e.g. motivation, attitudes, learning styles, self-efficacy - whilst student knowledge and comprehension can be more easily diagnosed through paper or online tests. ITSs also perform continuous observations of student behaviour linked to student performance in order to adapt to provide suitable feedback for encouraging interest and learning (Woolf 2009). Therefore, the student model is a representation of knowledge for the purposes of classification or prediction (Han and Kamber 2006).

Frequently, cognitive student models are derived from domain models or expert knowledge models, since these are composed of the concepts that students have to grasp and comprehend or techniques that students may use for solving a specific problem or case study. Therefore, domain models represent these facts and methods, which are signalled by expert lecturers as those required to solve domain problems, or those related to the instructional technique employed by these experts to provide feedback. The level of difficulty and time involved in the process of representing the domain is influenced by domain complexity or structure (Woolf 2009). Exemplifying this point, Anderson and Skwarecki (1986) created LISP tutor, which included hundreds of rules representing all possible tutor responses to student questions. The teaching strategy employed is known as model-tracing where students are constantly observed during problem solving tasks. LISP tutor provided instant responses to student actions, which characterised the need for early generations of ITSs to implement more advanced instructional strategies. In addition, attempting to represent all potential responses is difficult and highly time consuming. GUIDON, an ITS employed to teach medicine and engineering deployed Socratic Tutoring as its teaching strategy employing questions,

presented consecutively, to help students identify misconceptions and required knowledge. The domain was characterised with production rules, which provided educational feedback in the form of dialogues (Clancey and Buchanan 1982).

The earlier generations of ITSs had difficulty achieving an enhanced understanding of student needs (Du Boulay and Luckin 2001). As a result, instructional strategies also had problems with implementation and adaptation. This limitation was due to the AI techniques of that time, which could not characterise lecturer and student meta-processes. These limitations are not evident since the 1990s and 2000s where more machine learning representations and techniques are available, such as artificial neural networks (ANNs) and decision trees (Han and Kamber 2006, Russell and Norvig 2010).

Ohlsson (1987) relates key observations that have to be considered by ITSs to achieve successful learning. From these observations, those relevant to student modelling are: (1) given that the domain can be characterised in diverse ways, there may be several finish conditions for a specific case, situation or problem in a domain, (2) ITSs have to diagnose student knowledge and understanding as a key step in their teaching strategy, (3) student knowledge should be employed to select the most suitable pedagogical action and (4) the ITSs educational strategy has to consider the student characterisation and performance and learning goals and their close relation to pedagogical actions.

Also, there are models that distinguish between novice and expert student behaviour, which are easily derived from expert domain models. Whilst the latter have been characterised per task, these models are known as *overlay student models* (Woolf 2009). Noguez and Sucar (2005) employed an overlay student model to distinguish between novice, average or medium, and expert students whilst interacting with a VLE for teaching robotics. Student knowledge was represented and diagnosed over time using Bayesian Belief Networks (BBNs), which were derived by applying Probabilistic Relational Models (PRMs) with student observable behaviour.

There are also student models focused on characterising student misconceptions, which are usually derived from bug libraries comprised of student errors that are compiled continuously. Goguadze et al. (2011) investigated the implementation of cognitive and Bayesian student models, which could identify student errors while attempting to solve Maths problems. To achieve this goal, they employed data logs corresponding to 255 students at middle school level. The amount of data available for student modelling and its quality is known as *bandwidth* (Woolf 2009). Data was divided into training and testing sets, 70% and 30%. Goguadze et al. (2011) employed the most commonly observed student misconceptions and accordingly the most probable sources as a basis for their model. Therefore, each misconception is represented by a node with two possible states, present or absent. Finally the BN, characterising student misconceptions and their potential causes, was comprised of 12 mis-

conception nodes and 126 evidence nodes or possible answers to decimal problems in pre-tests and post-tests. Goguadze et al. (2011) evaluated their model with three metrics: (1) predicting student answers, (2) predicting correct student answers and (3) predicting erroneous student answers. The average accuracy of prediction for each metric was 60%, 69% and 87% respectively.

Open Learner Models (OLMs) are another type of student model. In this approach students are more involved in the model definition and refinement (Ahmad and Bull 2008). They enable students to reflect on their knowledge, skills and preferences and to share responsibility in their education. As a result, students become more independent. OLMs are initially derived from experts' viewpoints on the knowledge and understanding that these students must have. Students have access to OLMs with the purpose of evaluating the computer tutor knowledge and the meta-processes employed by the tutor to achieve this knowledge. Therefore, students can modify the computer tutor knowledge if this is incorrect. According to Ahmad and Bull (2008) limitations of this approach are that it is not really known to what degree students have an accurate perception of themselves nor to what extent students trust or believe in the computer tutor decisions, recommendations and the manner in which they use this information.

Mabbot and Bull (2006) implemented the Flexi-OLM system, which focused on representing students understanding about the C programming language. To derive their OLM they employed multiple-choice and short-answer questions. Their OLM was available in seven presentation formats in order to enable students to modify the model, influence the Flexi-OLM system to change it or negotiate these changes with Flexi-OLM. The seven ways in which concepts and their relations were characterised are hierarchy, logical grouping, arranged as lecturers from a course, concept map, index, a rank of concepts according to abilities and a summary. Mabbot and Bull (2006) conclude that students found it useful to have access to the learning model and they have preferences for selecting an editing presentation format. Also from the student viewpoint the Flexi-OLM system assesses students' knowledge and as a result the editing function was perceived as dishonest. Therefore, students preferred to negotiate their abilities with the Flexi-OLM system through a series of questions, which were presented to the students in order to ensure that they achieved the objectives associated with the specific concepts.

According to Woolf (2009) limitations of student models are:

1. Student models are qualitative, since they focus on characterising student behaviour. Student models comprise spatial, temporal and fundamental descriptions and relations of objects and methods.

2. Student models cannot completely account for student behaviour. Hence, student model accuracy is judged according to how useful the result is, instead of only taking into account fidelity of representation. It is observed that a more complete student model may result in more computational effort, which may not be justified.

Initially ITSs were focused on only assessing, understanding, diagnosing and representing student knowledge and subject domain knowledge as can be observed from LISP tutor (Anderson and Skwarecki 1986) and GUIDON (Clancey and Buchanan 1982). However, it was observed that one-to-one human tutoring is still more effective than computer tutoring (Du Boulay and Luckin 2001). Human tutors select suitable pedagogical strategies employing not only information about student knowledge and knowledge about the subject, but also other kinds of information, such as student preferences, confidence, effort, interest, motivation, attitudes, emotions, moods, self-efficacy, body language, voice tone and inflection and facial gestures (Alexander and Sarrafzadeh 2008). Hence, new challenges were set in order to enhance student awareness and a new generation of ITSs have arisen, which aims to be as successful as human tutoring (D'Mello et al. 2008, Sarrafzadeh et al. 2008). However, it is still not clear to what degree we will be capable of implementing these perception capabilities on ITSs (Du Boulay and Luckin 2001) or to what degree ITSs would and should behave as human tutors when providing instruction (Lepper et al. 1993).

A key goal of this new generation of ITSs is to recognise or understand emotion in order to respond to it appropriately. A motivator of this endeavour is the advent of the field of Affective Computing, which has influenced investigation in other fields (Picard 1995, Picard et al. 2004), e.g. ITSs and Serious Games (Sykes 2006). Another main motivator is work on ITSs that has focused on understanding and updating student disposition or motivation, which has recently proved to be closely related to student cognition and emotion (Pekrun et al. 2007, Porayska-Pomsta et al. 2008). As a result, the following section focuses on new generation ITSs.

3.3 A new generation of ITSs

ITSs are currently focused on achieving an enhanced understanding of student needs, concerns, expectations and preferences. This section examines these ITSs in order to identify how student modelling capabilities are achieved, i.e. the kinds of knowledge represented, approaches and AI techniques employed, evaluation techniques applied and existing challenges whilst attempting to reason about and recognise student disposition or motivation and affective states, i.e. moods and/or emotions

3.3.1 EDUCE

EDUCE is a web ITS that teaches Social Science and is implemented in Java and eXtensible Markup Language (XML). EDUCE focuses on identifying student individual characteristics and matching them with the appropriate instructional technique to achieve enhanced student engagement and motivation. Kelly and Tangney (2002) use Gardner's concept of multiple intelligence and learning goals defined at three levels for classifying student characteristics: levels facts memorisation, concepts understanding and problem solving.

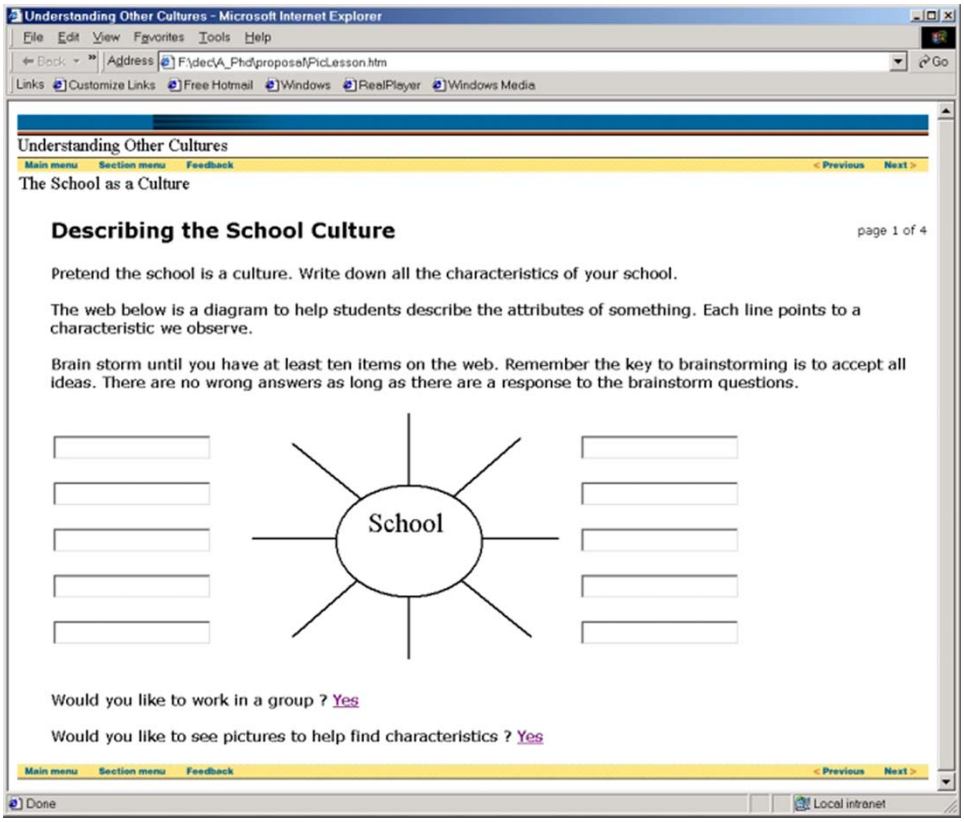


Figure 3.5 EDUCE GUI (Kelly and Tangney 2002, p. 736)

Whilst students interact with EDUCE, they choose material according to their preferences. All student actions are monitored by EDUCE and are employed to update the student profile. A report of the learning style identified is presented at the end of the interaction to each student. The goal is to assist students gaining self awareness of their own learning style. Figure 3.5 shows an example of the EDUCE GUI. Interaction variables are arranged as a two dimensional array that comprises all the possible combinations of student intelligence and learning goals. Other information recorded is a weighting factor that represents student learning preferences, navigation history, time spent in each learning unit, answers to questions and sections covered.

Gardner (1983) distinguished eight types of student intelligence that are possessed to different degrees, e.g. logical/mathematical, bodily/kinaesthetic, interpersonal and naturalist.

These are employed under different circumstances and interact in diverse ways. Hence, it can be inferred that according to the learning goal or goals – knowledge, comprehension, application, analysis, synthesis or evaluation - pursued by students, different types of intelligence is applied. Kelly and Tangney (2002) associated, through rules, each form of intelligence with available forms of interaction in EDUCE, which can be selected by students, e.g. naturalist intelligence utilises examples in nature in order to explain experiments. Every time that an item of interaction that represents a determined type of intelligence is selected, the count increases for that specific type of intelligence.

3.3.2 Intelligent Training System on the Internet

Souto et al. (2002) created an Intelligent Training System for real-life telecommunications settings on the internet as part of the Tapejara Project funded by the Brazilian Research Council. The Intelligent Training System is focused on assisting the instruction of employees in a Telecommunications company and its architecture is implemented through multiple agents. The student model agent identifies student cognitive learning styles with the Ross Test for Cognitive Processing, e.g. analytical and deductive reasoning, abstract relations and attribute analysis, and Bloom's Upper Cognitive Activities. The methodology employed by Souto et al. (2002) is centred on identifying the relevant cognitive learning styles and associating them to the different and available trajectories for solving a problem using correlations and factor analysis.

The employed observable features monitored by the student model agent are comprised of navigational, temporal and performance indexes, for example, the number of pages visited, mean time spent per web page and number of correct answers respectively. The AI technique employed for reasoning about these features and identifying the specific cognitive learning style is a Bayesian Belief Network (BBN). A log of interaction is maintained for every student. The Intelligent Training System GUI displays available content, which may be selected by the student for learning the topic of Time Division Multiple Access (TDMA).

3.3.3 MORE

Del Soldato and Du Boulay (1995) focused on modelling and handling motivational features related to the instructional process simultaneously to those derived from domain knowledge. The motivation for their research arose from observing that human teachers combine motivational strategies with domain-based strategies. Del Soldato and Du Boulay (1995) chose to model student psychological characteristics related to motivational strategies signalled by Keller (1983) and Malone and Lepper (1987), e.g. control, independence, confidence and effort. This motivational student model and the instructional-motivational planner were implemented in the MOtivational REactive plan (MORE) System (Del Soldato 1993). It is important

to signal that the instructional domain and motivational strategies sometimes complement each other, but sometimes are in contradiction according to Lepper et al. (1993).

Del Soldato and Du Boulay (1995) adapted the motivational strategies by Keller (1983) and Malone Lepper (1987) according to the context and capabilities of an ITS. In addition, the motivational features chosen had to be related to student observable patterns of interaction, e.g. effort was related to the time invested and the quality of the performance achieved. The methodology employed to identify the student level of motivation was first to use a questionnaire to assist students evaluating their level of confidence and motivation to study a specific domain, in this case teaching the Prolog programming language. Secondly, during students interaction with MORE communication was supported through a series of standard expressions and probable and possible answers to questions. The standard expressions and options in GUI menus were associated to students' possible levels of confidence, independence and effort, which are related to students' optimal state of motivation. For example, if a student asks very frequently for help it may be possible that the student has a low level of confidence or lacks independence. A final main step in this methodology comprised of asking and enabling the student to self-report regularly their motivational state while interacting with MORE. A limitation of the findings reported in Del Soldato (1993) is including participants or students (five Ph.D. students) to test and evaluate MORE who did not have previous experience using Prolog or were not particularly interested in learning it. Figure 3.6 shows an example of a displayed Prolog problem and MORE GUI. Figure 3.7 illustrates MORE's dialogue window to maintain communication with students.

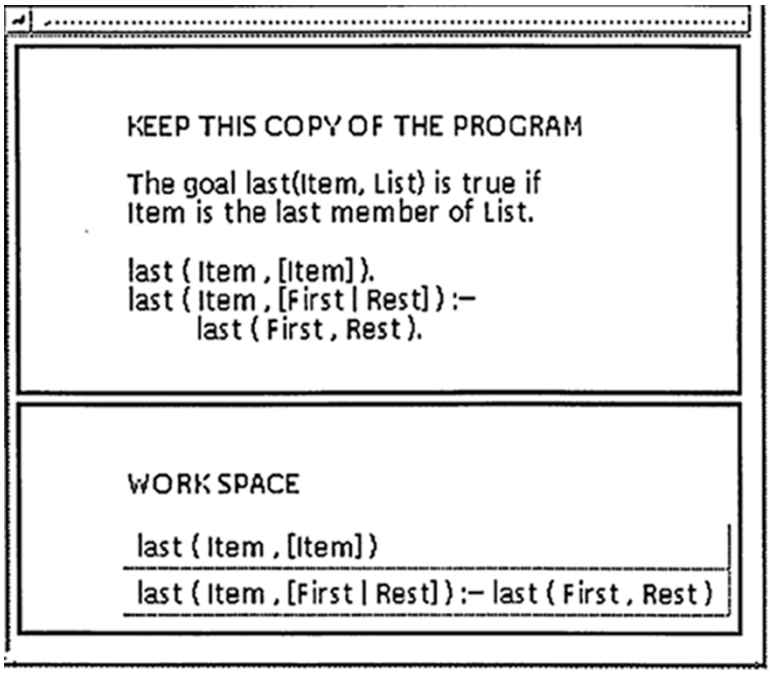


Figure 3.6 Example of Prolog problem and MORE GUI (Del Soldato 1993, p. 79)

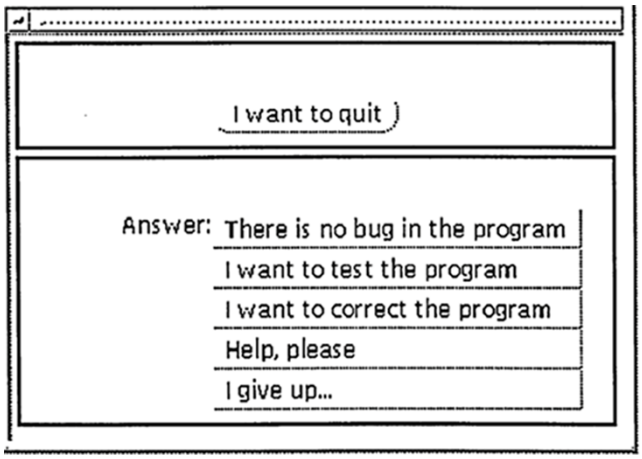


Figure 3.7 Dialogue Window in MORE (Del Soldato 1993, p. 80)

3.3.4 MOODS

De Vicente and Pain (2002) focused on the diagnosis or identification of student motivation and how student knowledge has to be formalised in order to achieve this goal. Being aware of the limitations faced by ITSs when compared to human tutors, their methodology comprised showing to a lecturer the previously recorded student interaction with an ITS in order to identify the relevant and probable observable variables and features related to student behaviour, e.g. mouse movements, that can be employed. The MOtivatiOn Diagnosis Study (MOODS) system is an ITS with characteristics of VLEs and GBL environments created by De Vicente and Pain (2002) to teach Japanese numbers. It comprises simple memorisation activities and game activities like Tetris. MOODS comprises capabilities to enable students to self-report in relation to their level of motivation though a series of sliders (De Vicente and Pain 1999). Figure 3.8 illustrates MOODS GUI. De Vicente and Pain (2002) employed the findings of Malone and Lepper (1987) and Keller (1983) in addition to the findings of Del Soldato and Du Boulay (1995). However, their conceptualisation of the knowledge within the student model was slightly different. For example, MOODS included what they defined as (1) permanent characteristics, which are comprised of student independence, attitude to challenge, control and expertise and (2) transient characteristics, which are comprised of student effort, confidence, sensory and cognitive interest, satisfaction and the relevance of the task to student goals.

MOODS was evaluated with 10 students at post-graduate level with or without previous teaching experience. The methodology employed included briefing the participants on the available student characteristics. Then participants watched the videos about students interacting with MOODS. Whilst watching these videos, participants were encouraged to infer students' motivational state and to provide comments about the possible sources influencing it. Participants also were enquired about the most suitable strategies to provide domain-

based and motivational instruction. Results were recorded on CD, transcribed and analysed. De Vicente and Pain (2002) collected in total 85 inference rules through this procedure. These rules required further validation in order to keep only the most coherent and appropriate set of inference rules (De Vicente 2003). Sixty-one inference rules comprised the final set, which provide more insight about the factors to be taken into account in order to diagnose student motivation.

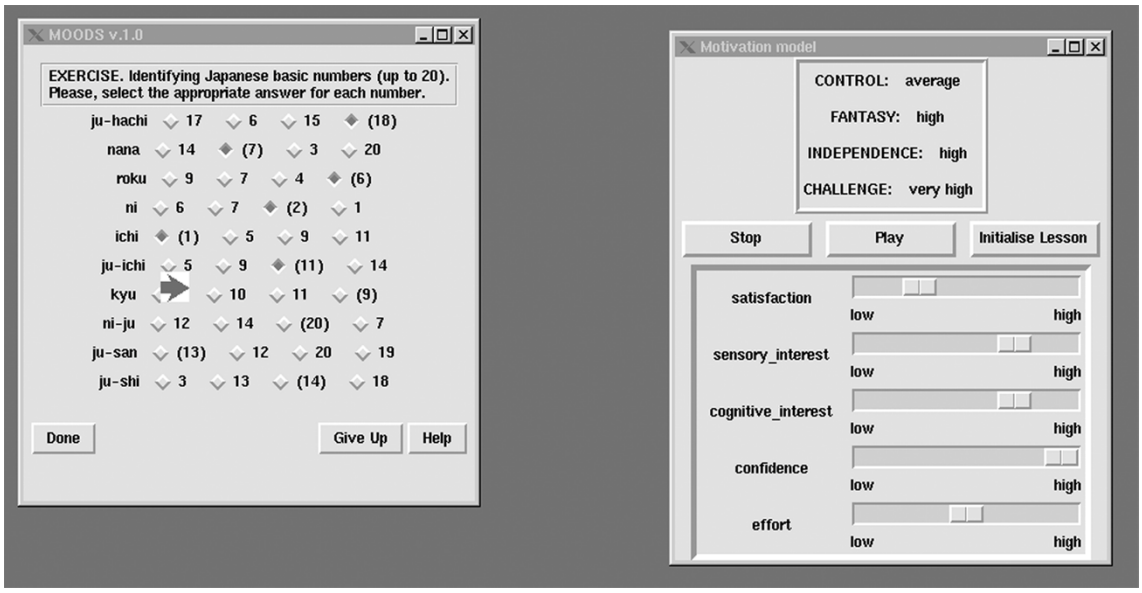


Figure 3.8 MOODS GUI (De Vicente and Pain 2002, p. 936)

3.3.5 M-Ecolab

M-Ecolab is a VLE that teaches Ecology to children between 9 and 11 years old. The main concepts discussed are related to food chains and food webs. The latter are graphical schemas showing relations among different species using as a reference the flow of energy. Rebolledo-Mendez et al. (2006) based their research mainly on the work of Del Soldato and Du Boulay (1995). Hence, they also focused on reasoning about students' meta-cognitive processes and variables related to student motivation, e.g. confidence, independence and effort. However, the key difference is the use of an Embodied Pedagogical Agent (EPA), known here as 'Paul', to provide guidance and instructional strategies. M-Ecolab is an enhanced version of the Ecolab VLE created by Luckin and Du Boulay (1999), which only focused on domain-base instructional strategies. Figure 3.9 shows the Ecolab GUI. Rebolledo-Mendez et al. (2006) evaluated M-Ecolab motivational support with 19 children, 10 boys and 9 girls.

The methodology employed comprised a pre-test in food chains and webs and another pre-test to measure students' initial motivation. The latter was based on Harter's test employed in primary schools in the UK. Students learned to use M-Ecolab through a video tutorial. Afterwards students interacted with M-Ecolab in two sessions for 40 minutes. After inter-

acting the second time with M-Ecolab, students filled in a post-test. Students were categorised in three skill groups, low, average and high skilled based on the Standard Assessment Test (SAT). Finally, four groups were identified during investigation: (1) motivated students before and during their interaction with M-Ecolab, (2) motivated students before and with low motivation during the interaction, (3) unmotivated students before and with high motivation during the interaction and (4) unmotivated students before and during their interaction with M-Ecolab. The learning gain of the last two groups of students was significantly greater, 40.90% and 27.27% respectively, in comparison to the first two groups of students, 12.87% and 17.42% respectively. Analysing the motivational variables effort, independence for the groups with the major learning gain, groups 3 and 4, it was observed that they have significant differences in the effort invested and the exercised independence, but not in their level of confidence. Also, it was noted that students in group 3 followed and paid more attention to the suggestions of the Paul M-Ecolab EPA. These, results suggest that motivation techniques should be deployed with care when used with students that are already motivated.

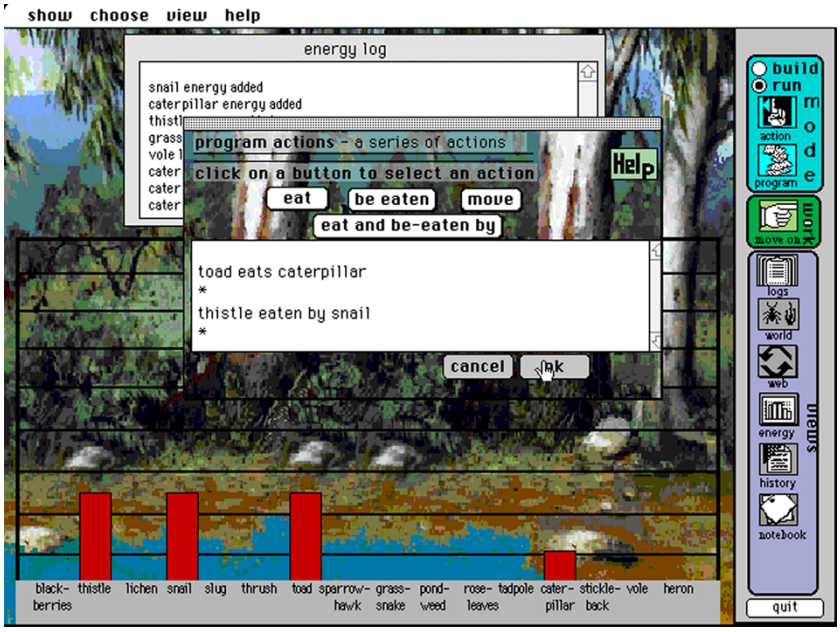


Figure 3.9 Ecolab GUI (Luckin and Du Boulay 1999, p. 202)

3.3.6 Wayang Outpost

Arroyo and Woolf (2005) focus on identifying student goals and attitudes using interaction behaviour stored in data logs. Their ultimate goal is to encourage and keep a positive attitude towards learning. A key problem is determining dependencies between observable behaviour variables and variables corresponding to students' psyche. Arroyo and Woolf (2005) employed the data logs from student interaction with Wayang Outpost, an online ITS, developed

to teach Maths to students between 15 and 17 years old at high school level. An example of the Wayan Outpost GUI is shown in Figure 3.10.

Wayang Outpost instruction and feedback comprise videos with audio that guides students step by step in solving problems and assist students with learning concepts. Examples of interaction variables monitored in data logs were reported: the time spent solving problems, the problems selected and the interval of time taken to give a response. A data log was derived from the interaction of 230 students with Wayang Outpost, which took on average between 2 to 3 hours. On completion of student interaction, a post-test of the topics presented by Wayang Post was then taken by students and they also completed a survey regarding their attitudes.

From analysing the acquired data logs, Arroyo and Woolf (Arroyo and Woolf 2005) distinguish between different kinds of student interaction. It was observed that these forms of interaction reflect student effort, attention and problem solving strategies. In addition, it was noted that student behaviour related to searching and requesting help can be categorised. Other features distinguished are past-experience derived from observed student performance solving a pre-test and their gender. To determine dependencies between student attitudes to learning and the interaction variables, Arroyo and Woolf (2005) considered the relevant Pearson correlations. A Pearson correlation is a measure of a supported linear relationship between two variables (Kinnear and Gray 2010). For example, the learning gain had a positive correlation with the time spent receiving help with the attempted problems. The identified correlations did not prove significantly strong, i.e. by themselves they cannot account for more than 15% of the variance. However, the integration of these variables may allow an enhanced prediction of student attitudes related to their learning experience.

After identifying the relevant correlations, they are then employed to create a Bayesian Belief Network (BBN) comprising the interaction variables and the variables related to student psyche. A Chi-Square test was employed to eliminate links to some variables in the DBN and to determine which dependencies were not preserved after converting them to discrete or categorical variables. Finally, the Conditional Probability Tables (CPTs) are set from cross-tabulations derived from student data, e.g. maximum likelihood. Arroyo and Woolf (2005) evaluated the derived Bayesian student model through conducting a 10 fold cross-validation where 90% of student data was employed for training, i.e. learning CPTs, and 10% was employed for testing. Evidence was propagated in all the leaf nodes of the DBN corresponding to the interaction observable variables, e.g. pre-test mark and gender. Through the propagation hidden nodes, student attitudes and other cognitive variables, such as perception of helpfulness, were inferred. The inference is considered uncertain if it falls in a range between 0.3 and 0.7, i.e. the inference is not fired. Only certain inferences were taken into account for the evaluation. A key finding of this investigation was to create a student model that clearly

depicts the relationships between student attitudes and selected variables and, as a result, facilitating its human understanding.

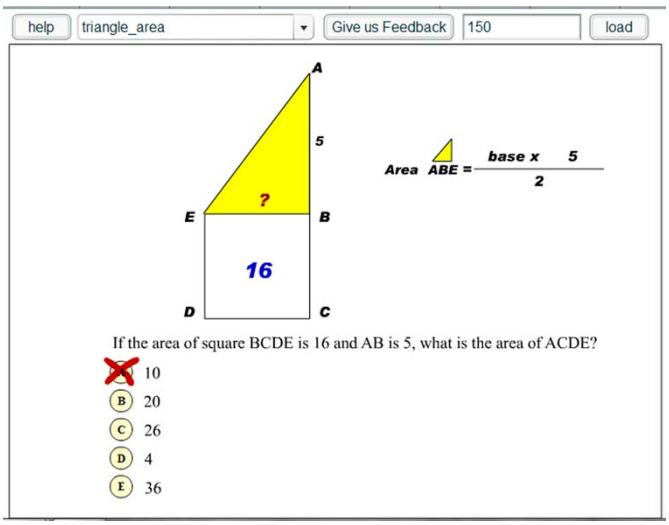


Figure 3.10 Wayang Outspost GUI (Barney 2011)

When Wayang Outpost was implemented in educational settings, researchers created 25 packages of sensors to bring to classrooms. Each package comprised: a chair with sensitive cushions and accelerometers, a GSR bracelet, facial recognition camera and pressure mouse. Half of the data acquired through the sensors was not complete owing to a technical failure. Therefore, this technology does not provide a very practical solution. Also, it requires intensive processing and the information and signals acquired through sensors, which require high bandwidth.

3.3.7 CRYSTAL ISLAND learning environment

McQuiggan et al. (2008) focused on identifying student level of self-efficacy, which is related to student beliefs about performing successfully in specific situations and intertwined with student affective states. To achieve this goal, McQuiggan et al. (2008) endeavoured to create a student model that can diagnose student self-efficacy and can notify pedagogical decisions. It is important to note that self-efficacy is domain-specific, e.g. students who have a high level of self-efficacy in Maths may have a low level of self-efficacy in History. Static and dynamic student models were created, where the former was derived from data of an applied pre-test and the latter is derived from the same pre-test data, but also from student physiological data - e.g. heart rate and galvanic skin response (GSR) – and the observation of student interaction in CRYSTAL ISLAND a narrative-centred learning environment for teaching Microbiology and Genetics to students at high school level.

The CRYSTAL ISLAND GUI is shown in Figure 3.11, where students play a detective who needs to solve a mystery. CRYSTAL ISLAND was developed with the source engine and 3D game platform by Valve Software. To create a self-efficacy student model, McQuiggan et al. (2008) observed that this needs to be implemented with computational and AI techniques that operate at runtime, must support and satisfy demands of real-time and interactive learning and may not disrupt the learning process. For deriving and evaluating their self-efficacy model, McQuiggan et al. (2008) used data logs, answers to the problem solving self-efficacy instrument by Bandura (2006), and physiological and demographic data corresponding to 33 undergraduate students with an average age of 26 that interacted with CRYSTAL ISLAND for learning Genetics. This may be a limitation for this investigation, since the application targets students at high school level, but was designed with students at undergraduate level, whose level of self-efficacy may be influenced by their age. Physiological signals are monitored and captured whilst students interact and solve problems in CRYSTAL ISLAND. After solving each problem students self-report their level of self-efficacy through a slider. The architecture of CRYSTAL ISLAND is SELF, which is designed to interact with students to acquire data and to monitor student level of self-efficacy through real-time observable behaviour.

An inductive approach was employed to derive naive Bayes classifiers and decision tree self-efficacy student models. McQuiggan et al. (2008) used as a basis the work of Bandura (1997, 2006), where four kinds of sources may influence student levels of self-efficacy: (1) enactive mastery experiences, (2) vicarious experiences, (3) verbal experiences and (4) physiological and emotional effects. Additionally, self-efficacy also influences student cognitive, motivational, selective and affective processes. Observable and available features at runtime are employed to derive the self-efficacy student model, e.g. temporal features (time spend on each question), location features (position of mouse), intentional features (number of goals achieved) and physiological features (GSR and heart rate). A vector of 144 interaction and observational variables was recorded and employed to derive the naive Bayes and decision tree models, where machine learning techniques were employed to select the relevant features. A 10 fold cross-validation procedure was employed to evaluate the model's accuracy. Results showed that static and dynamic models are significantly accurate, effective and support learning. The naive Bayes classifier achieved 82.1% accuracy whilst the decision tree model achieved 87.3% accuracy. It was observed that when incorporating physiological variables, average heart rate and average GSR, the model increased its accuracy by only 10%.

Sabourin et al. (2011) employed CRYSTAL ISLAND in order to predict the valence of mastery and learning goals, e.g. positive or negative, and the emotions related to this valence. To achieve this goal they derived an affective student model based on the appraisal-based the-

ory of learning emotions by Elliot and Pekrun (2007). This theory considers that a student has two approaches to achieving a goal: (1) performance and (2) mastery. Students that pursue performance seek competence and competition with other students whilst students that are focused on mastery seek to achieve skills. The outcome of both goals is evaluated as succeeded or failed, i.e. positive or negative valences. The valence is a characteristic of achievement emotions, which are emotions relevant to the learning context. Some students can be more focused on the positive valence of the goal than the negative one and vice versa. This theory links the achievement of goals and its valence with the experience of achievement emotions in a learning context.

Figure 3.11 CRYSTAL ISLAND GUI (McQuiggan et al. 2008, p. 108)

The methodology employed by Sabourin et al. (2011) comprised using CRYSTAL ISLAND in order to teach Microbiology to 296 students at high school level. CRYSTAL ISLAND is a role playing game in which the player has to discover the source of a disease and the kind of disease that it is spreading. The student has to interview and interact with NPCs in order to solve the mystery, e.g. handling objects or lab equipment and making notes. The student is allowed to perform free exploration and the game does not have a clearly defined problem-space. Therefore, game goals can be perceived in some degree as blurred. While reviewing the student interaction data acquired to design and validate their model, they noted logging errors and incomplete data. Hence, the authors rejected some data leaving them with data from 260 students.

Before interacting with CRYSTAL ISLAND, pre-study material was completed. In addition to a demographic survey, students completed a Big Five personality questionnaire, a goal orientation questionnaire and an “affect regulation tendencies” questionnaire. All this information was employed in conjunction with the data collected during student interaction to evaluate the model. The structure of the DBN employed is shown in Figure 3.12. As can be ob-

served the valence and the type of focus that pursues the student are the main antecedents of the emotional state experienced. During game interaction students reported their emotional state through a smart-phone device. The emotions that were investigated during this study were anxiety, boredom, confusion, curiosity, excitement, frustration and concentrated. Sabourin et al. (2011) tested the performance of their model implemented with Naive Bayes, BBNs and DBNs. The DBNs model was the most accurate when predicting emotion, with 36% accuracy, and valence, with 72.6% accuracy. In their model they did not include no-emotion or neutral emotion, since these are not included in the model by Elliot and Pekrun (2007).

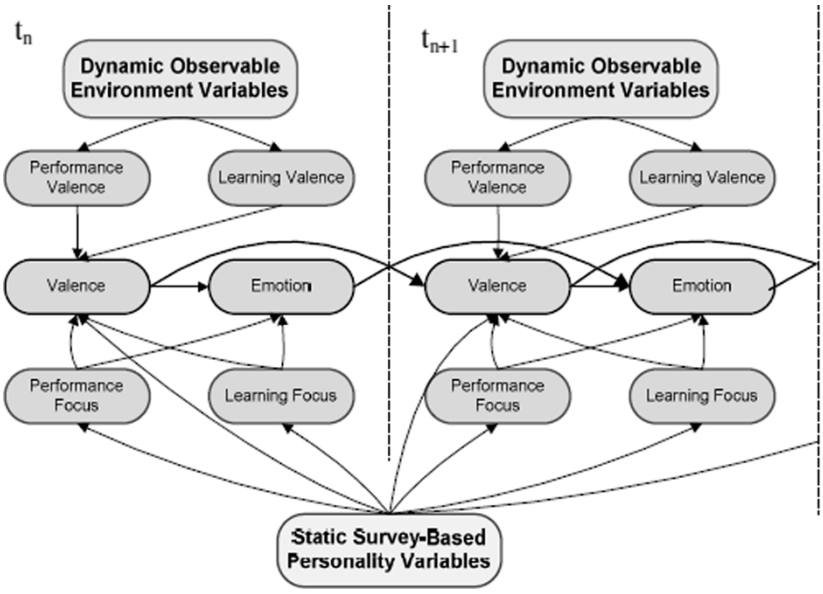


Figure 3.12 DBN based on appraisal-based theory (Sabourin et al. 2011, p. 292)

3.3.8 AutoTutor

D'Mello et al. (2008b) focused on recognising relevant affective states related to the learning process in real-time through monitoring body language, facial gestures and features in dialogues. These features are acquired through hardware equipment, e.g. sensors and cameras, and are mapped to affective states determined by expert judges. As an example, body language was monitored though the Body Posture Measurement System (BPMS) and facial gestures through the IBM BlueEyes system. Also, written natural language was analysed though Latent Semantics Analysis (LSA) (D'Mello et al. 2008a), which is employed to analyse statistically, relationships between documents that are part of a set and their terms. A limitation of using sophisticated hardware equipment is that is not easily accessible anytime-anywhere by final users. In addition, to interact online with such systems, participants have to work with a limited set of capabilities, e.g. written natural language. Therefore, full access to the capabilities of such systems is only achieved in physical laboratories built for this purpose

(Burleson and Picard 2007). However, additional sensors and hardware equipment is prone to failure and expensive.

This affective student model derived from student physical patterns by D'Mello et al. (2008b) is included in AutoTutor, which is a combination of a VLE and an ITS implemented for teaching students at undergraduate level about critical thinking, computer literacy and Newtonian physics. Physics has been adopted as a topic to develop more interactive teaching tools, since it has a reputation for being 'difficult' and 'boring' (Sillitto and MacKinnon 2000). An example of AutoTutor's GUI is shown in Figure 3.13. As can be observed, students can interact through text and AutoTutor can evaluate student written explanations, provide hints, summarise and correct misconceptions (D'Mello et al. 2008a). An EPA is also employed to provide affective feedback. Training and testing data corresponding to 28 participants was acquired. D'Mello et al. (2008b) mapped student physical effects to previously identified and relevant affective states in the educational context, e.g. boredom, frustration, confusion and engagement. Also, they endeavoured to change negative affective states to one that encourages flow, i.e. engagement.

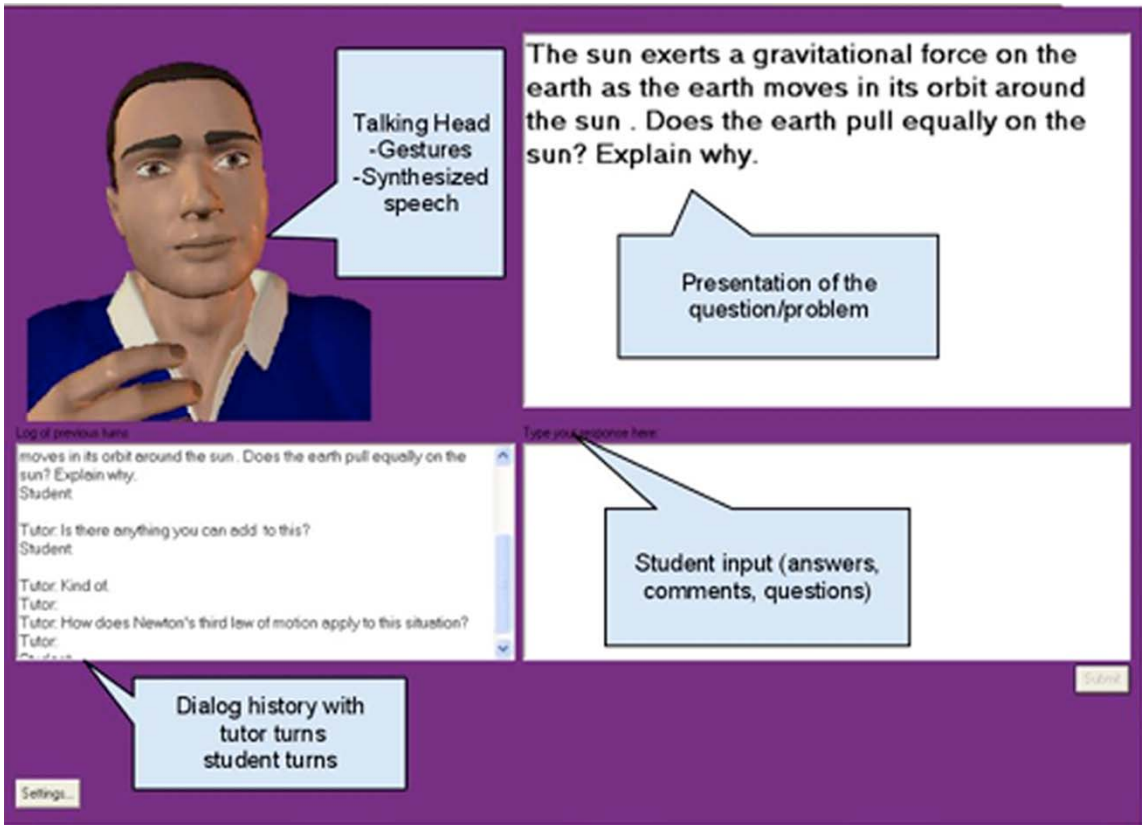


Figure 3.13 AutoTutor GUI (Henry 2011)

During evaluation of the identification of patterns, the BlueEyes system showed 65% accuracy. To link facial gestures to emotions in addition to the opinions of expert judges, the facial

coding system by Ekman and Friesen (1978) was employed as a basis, who observed that across cultures there are universal gestures shared by individuals to communicate affective states. Seventeen classification algorithms that use cross-validation, such as Bayesian classifiers, ANNs with supervised learning, functions, trees or rules, were applied for indentifying conversational features achieving 69%, 68%, 71% and 78% accuracy for boredom, confusion, engagement and frustration respectively. Also, these algorithms and AI Techniques were applied for classifying body language achieving 70%, 65%, 74% and 72% accuracy for boredom, confusion, engagement and frustration respectively. From a high level viewpoint, AutoTutor classifies with 73% accuracy the relevant affective states.

3.3.9 Prime Climb

Conati (2002) and Conati and Maclaren (2009) endeavoured to create an emotional student model that can identify student emotions in real-time through observable behaviour, physical and physiological effects of emotions. The affective student model is implemented with Dynamic Bayesian Networks (DBNs). This approach employs the Ortony, Clore and Collins (OCC) model, which explains how emotions arise from a cognitive viewpoint and determines various types of emotions according to their source (Ortony et al. 1990). According to this theory, sources that can elicit an emotion are events, agents and objects, which are evaluated or judged against individuals' internal constructs, e.g. events are evaluated against goals, actions against social standards and objects against attitudes. Determining students' internal constructs is a key problem with this approach.

To overcome this limitation, Conati and Maclaren (2009) assumed that internal constructs are influenced by personality traits. Hence, they employed the Big Five theory (Costa and McCrae 1992), which links personality traits with individuals' tendencies to experience affective states. Prime Climb, a GBL environment for teaching Mathematics, e.g. factorisation of numbers, by the Electronic Games for Education in Mathematics and Science (EGEMS) group comprises this affective student model as part of its architecture. Prime Climb is directed to students at primary level in an age range between 10 and 12 years old. Physical and physiological effects of emotion are mapped to emotional states by student self-reporting. Figure 3.14 shows the Prime Climb GUI and self-report dialog box. Students play against each-other in order to climb as quickly as they can to the top of the mountain. To attain this goal they must select correctly the positions in which the numbers are not factors of the number in which his or her opponent is located. The ultimate goal is to enable an EPA achieving suitable equilibrium of emotional feedback and pedagogical actions.

The methodology employed to infer student goals from observable interaction behaviour involved asking students to respond to a personality test and then asking them to interact

with Prime Climb through a Wizard-of-Oz technique (Andersson et al. 2002). This entails convincing students they are interacting with a computational system, when actually they are interacting with a human. Finally, students are asked to list the goals that they have when playing Prime Climb. Afterwards, data logs are employed to derive a DBN, which comprises relevant student goals, e.g. having fun and avoiding failure, and observable variables according to student personality type. Dependencies in the DBN were set with identifiable correlations. It is important to note that student goals can change over time and personality does not change at the same rate as student attitudes, which are dynamic features of personality, but are more related to student opinions about objects according to the information that they have from them (Ajzen 2005).

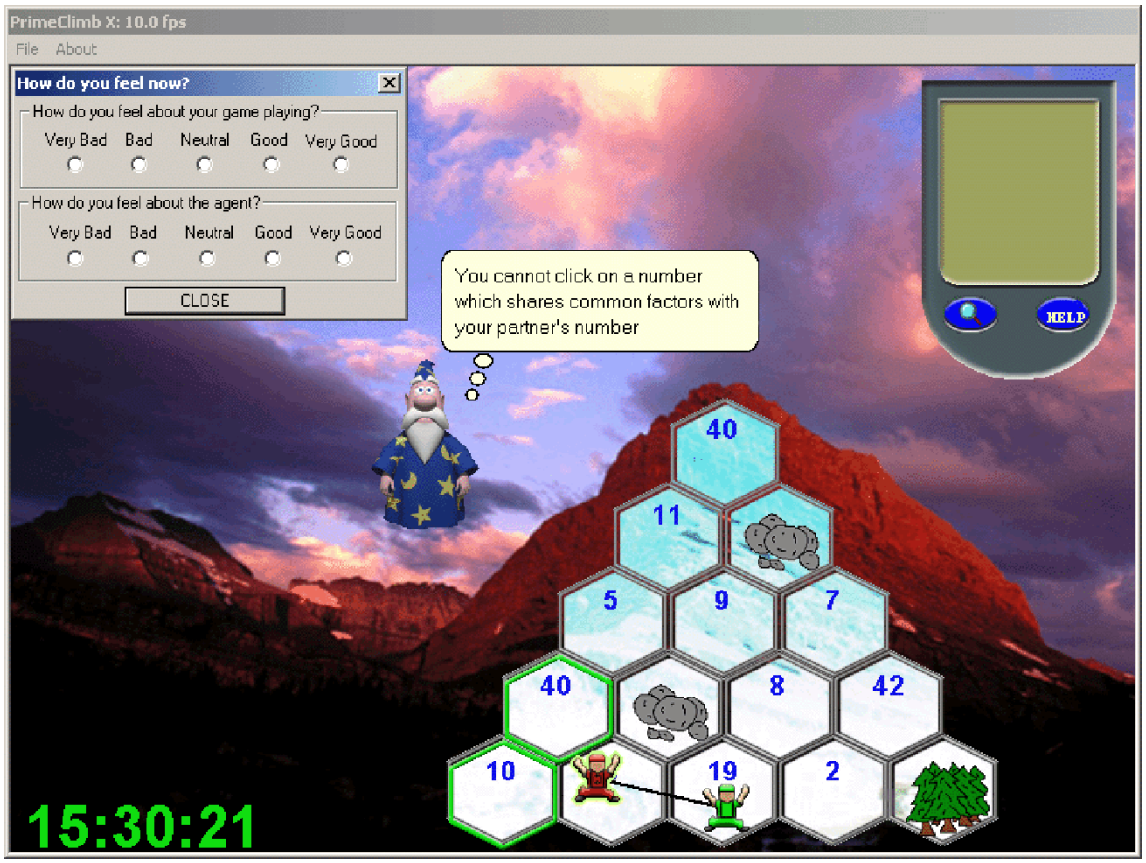


Figure 3.14 Prime Climb GUI (Conati and Maclaren 2009, p. 281)

From the 22 emotions defined in the OCC model, Conati (2004) selected the relevant emotions to the learning experience through a series of previous pilot studies. As a result, the emotional student model of Prime Climb focuses on distinguishing *joy and distress*, which students may feel towards the activity outcome, *pride and shame, which* can be experienced as a result of students' own actions and *admiration and reproach* as a result of EPA's decisions and instructional strategies. Self-reports of undergraduate students that interacted with Prime Climb were employed to derive, calibrate and evaluate the emotional student model by

Conati and Maclaren (2009). Two approaches were employed to evaluate the classification accuracy of the emotional student model: (1) micro-average, which involves knowing the percentage of cases correctly classified over all the test instances independently of their class, and (2) macro-average, which corresponds to knowing the percentage of cases correctly classified per class. As a result, Conati and Maclaren (Conati and Maclaren 2009) achieved 69.59%, 62.30%, 67.42% and 38.66% accuracy for the emotions joy, distress, admiration and reproach respectively.

3.3.10 Easy with Eve

Sarrafzadeh et al. (2008) also focused on recognising students' affective states in real-time whilst teaching Mathematics to primary students, specifically the topic of part-whole addition. To achieve this goal, Sarrafzadeh et al. (2008) recorded and observed one-to-one tutoring interactions acquiring insight about and data from lecturer and student behaviour, facial expressions and relevant affective states, e.g. surprise, laughter, smiling, fear, disgust and neutral (Alexander and Sarrafzadeh 2008). As an example it can be mentioned that the most frequent teacher action is to respond to student questions and usually they provide immediate positive and neutral affective feedback. Teachers rarely communicate negative feedback.

The EPA, Virtual Eve, is shown in Figure 3.15, has pedagogical strategies including showing empathy and strategies for creating a sense of affinity. Each user interaction was coded and labelled. Each case recorded was employed to select a suitable response to student identifiable needs. However, if the case is not matched, the query will be narrowed until finding a suitable match. As a result, Virtual Eve is not capable of responding to all student needs and questions. In addition, Virtual Eve has a text-to-speech generation module to respond to student questions.

Figure 3.15 Virtual Eve EPA (Alexander et al. 2006, p. 43)

Students' facial gestures were acquired employing a web-camera and mapped and classified with Ekman and Friesen (1978) research for mapping facial muscles and ANNs with supervised learning methods, such as Supported Vector Machine (SVM). To evaluate and train the student model a 5 fold cross-validation procedure was employed. Sarrafzadeh et al. (2008) achieved 94%, 96%, 93%, 90% and 93% accuracy while recognising surprise, laughter, smiling, fear, disgust and neutral. An embedded system was designed and implemented for integrating the facial and gesture recognition systems with the main aim of enhancing the overall performance of the system.

3.3.11 EMASPEL

Neji and Ben Ammar (2007) have the goal of recognising student affective states through analysing their facial gestures. Their aim is to investigate the influence of emotion on group and individual interactions. An approach comprising the temporal analysis of distances between static facial features, e.g. eyes, eyebrows and mouth was applied. The affective states attempted to be recognised by Neji and Ben Ammar (2007) are joy, sadness, anger, fear, disgust and surprise, which are the six basic and universal emotions defined in the research of Ekman (Ekman 1999). In addition to these emotions, Neji and Ben Ammar (2007) also focused on recognising the facial expression corresponding to a neutral affective state. The research by Ekman (1999) was also employed as a basis for mapping facial displays, represented as a pattern of distances, with identifiable affective states. Neji and Ben Ammar (2007) do not provide information about the accuracy of the approach or the AI or machine learning techniques employed to classify and map facial gestures to affective states.

This emotional student model and approach was implemented in the Emotional Multi-Agents System for Peer to peer E-learning (EMASPEL), which teaches Communications Technology to students at undergraduate level. As its name suggests, the EMASPEL architecture is comprised of multiple agents that perform and serve several functions in a coordinated manner, such as the *interface agent*, the main goal of which is to transmit the information on student affective state to other agents, such as the *tutoring agent*, and send the information related to student actions to the *agent curriculum*. EMASPEL employs an EPA to provide instruction and address affect. EMASPEL is implemented using the MadKit platform, which was developed using Java. Until now, EMASPEL's EPA focuses on mirroring students emotional state. It is important to note that knowing what is the most appropriate response to each of the different student affective states is the key activity.

3.3.12 ERPA

Chalfoun et al. (2006) focused on reasoning about student emotional states in an online learning environment, known as the Emotional Response Predictor Agent (ERPA). Their approach entailed being aware of student personality and events, which occurred during student interaction. Chalfoun et al. (2006) use the ID3 algorithm and the OCC model for creating their emotional student model. They focused mainly on predicting the emotional state of the student when knowing their final score in a test. Towards this goal, student beliefs about the prospective score was enquired and the obtained score was recorded and showed to students once they finished answering the test. Additionally, information about student attributes, e.g. sex and personality type, were also kept and employed to predict student emotion.

The investigation methodology employed by Chalfon et al. (2006) comprised registering and logging in to the system, answering a personality test based on Eysenck Personality Questionnaire (EPQR-A) (Francis et al. 1992) and also answering a quiz. The Quiz comprised items related to general knowledge such as Sports and Emotional Intelligence. The EPQR-A measures four dimensions: extraversion, psychoticism, neuroticism and lie scale. The latter is related to individuals' preferences for providing socially desirable responses. The predominant personality trait will be the one with the highest score. At the end of the quiz students are asked to report the expected score and are presented with the actual score. From the 22 emotions defined in the OCC model, Chalfoun et al. (2006) selected those that they considered relevant to the context and situation: disappointment, distress, joy, relief, satisfaction or fear.

The ID3 algorithm, run with 10 fold cross validation, was employed to derived classification models and decision trees from data. 139 individuals participated in the derivation and calibration of the emotional student model. 59 scored highest in the personality trait extraversion, 30 in neuroticism, 12 in psychoticism and 38 in lie scale. The architecture of ERPA, shown in Figure 3.16, comprises a number of modules. The rules extraction module has the function of analysing the ID3 structure (which has a tree shape) in order to adapt the values of the leaves. The latter will be used to extract the rules for training ERPA. After deriving and calibrating the student model, the model was included in ERPA and 34 online participants assisted to evaluate the model accuracy. Results showed that the model achieved 84% accuracy.

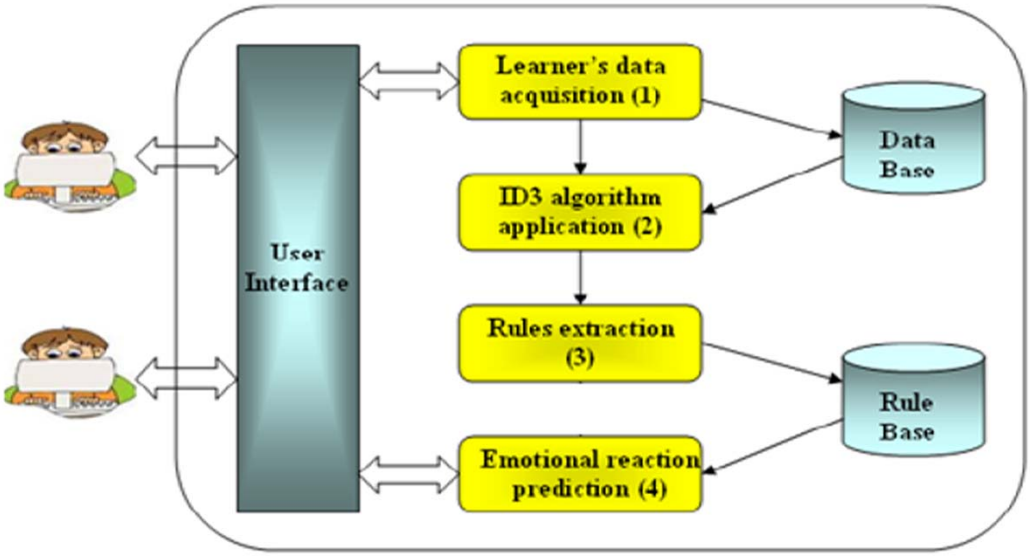


Figure 3.16 Architecture of ERPA (Chalfoun et al. 2006, p. 15)

3.3.13 The Collaborative Educational Environment and JADE

Jaques and Vicari (2007) focused on diagnosing students emotional state with the Belief-Desire-Intention (BDI) model by Georgeff et al. (1999) and OCC. The emotions to be recognised are satisfaction, disappointment, joy, distress, gratitude, anger, pride and shame. Their ultimate goal is to employ an EPA to encourage positive emotions using supporting messages and communicating affective strategies through attitudes. This emotional student model reasons about events and student goals and actions. It was included in the collaborative educational environment by Andrade et al. (2001).

The instructional approach of the collaborative educational environment is based on Vygotsky's socio-cultural theory (Vygotsky 1998). Vygotsky's theory explains that intellectual activities happen in a context where an individual establishes a relationship of a social kind with his or her environment through time. This relationship is regulated for what Vygotsky defines as instruments and signs. Signs are incentives or mechanisms for adaptation and they are used to solve specific problems and controlled internally, while instruments have a mediation function betwen the student and the goal of his work, e.g. asking for help while performing an activity, but controlled externally. Also Vygotsky mentions the Zone of Proximal Development (ZPD) that may be defined as the distance between the actual capacities of a student to independently solve a problem and the probable development that the same student can achieve with another individual or other individuals collaboration. In addition Vygotsky argues that cognitive and psychological processes occurred initially at a social level and afterwards at an individual level.

The collaborative educational environment by Andrade et al. (2001) is based on the previous concepts and ideas by Vygotsky and it was created as a multi-agent systems architecture, e.g. the ZPD agents are focused on identifying actual student skills and knowledge and suggest activities that encourage the development of required capabilities. Jaques and Vicari (2007) designed the student model considering that students can pursue two goals: achieving mastery or only an satisfactory performance. They considered that a manner of being aware of them is to enquire about student motivation.

After identifying student goals, events and actions related to the goals are classified by expert psychologists or teachers. Events can impede or facilitate the achievement of goals and this characteristic determines if events are desirable from the student perspective. It was noted that effort is related to student desirability of goal achievement. Hence, they employed the representation of student effort given by Del Soldato and Du Boulay (1995). To measure student motivation, a self-report test, the Motivated Strategies for Learning Questionnaire (MSLQ), was applied. Student emotion is determined through inference rules. Initially, the student model and the BDI approach was implemented in Prolog. Jaques and Vicari (2007) did not present results that show this model's accuracy. The BDI research approach considers that an agent is an intentional system and can communicate attitudes just as a human (Jaques and Vicari 2007). *Beliefs* are related to information about the environment. Desires are related to student motivation, e.g. goals and priorities. Intention is related to the student's goal, which is approached through a plan.

Jaques et al. (2011) discuss the evaluation of their emotional model. However, instead of employing the collaboration learning environment in their investigation, they employed JADE, an online learning environment that does not include an ITS. They asked students to self-report their emotions with dialog-boxes in a similar way to Conati and Maclaren (2009). They use one static dialog-box and one dynamic. In the self-report dialog-boxes Jaques et al. (2011) joined the emotions disappointment/distress and joy/satisfaction. This decision was taken after the observation that similar events in the system influenced similar emotions. Afterwards the emotion inferred by the emotional student model was compared with the emotion reported by the student. Jaques et al. (2011) decided to present the dynamic dialog box at a fixed interval of time or every time that the system detected an emotional change. They inferred that students did not perceive the process as invasive, since they did not report frustration frequently.

Twenty-four students aged between 12 and 19 at high school level in Brazil participated in this investigation. The agent Pat, implemented with a BDI approach, was included in JADE, and used in this investigation to teach about Earth time zones. Eight of the 24 students did not report any emotion. Researchers discarded the data, since they inferred that they did not want to reveal their emotional state. 8% of students reported the discomfort felt while report-

ing emotions in a post-questionnaire. Results showed that the group joy/satisfaction was classified with 77% accuracy. The emotions anger and disappointment/distress were not recognised at all and gratitude was recognised with 22% accuracy. Jaques et al. (2011) concluded that these results are owing to biased student self-reports, where students often reported gratitude toward the assistance of the agent.

3.3.14 LeActiveMath or LeAM

Porayska-Pomsta et al. (2008) focused on observing actual one-to-one instruction interactions to identify the affective states considered relevant by expert teachers and how these teachers respond to them. Data logs were recorded in order to identify student and lecturer actions. Machine learning techniques were employed in order to diagnose and predict student interest, confidence and effort. The learning environment employed on this investigation is the 'Language Enhanced, User Adaptive, Interactive e-Learning for Mathematics' system (LeActiveMath or LeAM) (European Comission 2008). LeAM teaches Mathematics, e.g. calculus, to students at high school and undergraduate levels. LeAM is an online system comprised of several components: an OLM, tutorial, domain reasoner and natural language modules and a repository of exercises. Communication through natural language is conducted through a chat interface.

Porayska-Pomsta et al. (2008) follow an observational approach known as the Persistent Collaboration Methodology (PCM), where data and dialog logs, verbal protocols, semi-structured interviews and post-task walkthroughs are triangulated to enhance inferences reliability. Observational variables feasible for low bandwidth were employed in this investigation, e.g. the interval of time taken by the student to give a response. The ultimate goal is to derive a series of rules to adapt feedback using an EPA that employs the theory of linguistic politeness by Brown and Levinson (1987) to provide instruction. The theory of linguistic politeness entails that the lecturer or teacher chooses and uses a level of indirect feedback according to the level of politeness that lecturers consider necessary in specific situations. Usually feedback is adapted according to student gestures or other psycho-physiological features. Politeness is expressed through language.

Porayska-Pomsta et al. (2008) employ tutors with at least 2 years of experience, since it has been observed that expert tutors assist students in achieving enhanced learning outcomes in comparison with novice tutors. This is the result of providing regular and effective feedback and having an enhanced awareness of student cognitive and affective needs and problems. In order to observe and analyse one-to-one teaching interactions, LeAM and the design of the study were pilot tested and then modifications were performed in both. The modified and updated design study was applied and finally more data was acquired. Five

participants were involved in the study, four at undergraduate level from Edinburgh University and one at high school level.

Participants interacted with students through exercises in order to teach them the chain rule, a topic from the subject of differential calculus that is taught during first year undergraduate level. Two interfaces were included and developed, one for lecturers and one for students. In addition, functionally was added to record and video capture student screens. The student GUI is comprised of a *theory frame*, in which concepts and background theory of the exercises are referred to and discussed, and *a text and math editor* where formulas can be written and explained. Students can then submit their responses to lecturers with the 'send' button. A window showing the *History of interaction* between the student and the lecturer is also included. In contrast, the lecturer's GUI is comprised of a frame, *the situational factors selection tool*, which assists them to annotate and identify features, variables and values that they considered relevant to provide their feedback

The lecturer GUI includes *a preview frame* to take a glance at the feedback that will be displayed in the student GUI and an exercise frame where pre-defined chain-rule case studies were provided in order to be used during the session. A second screen is employed by the tutor to monitor all students' actions anytime during the interaction. Twenty-six interactions were recorded in total. Data was analysed through identifying the frequency in which factors were employed by lecturers to diagnose student affective state. Correlations between factors and categories were obtained. Qualitative analysis of interviews and walkthroughs was conducted in order to identify other sources of evidence employed by lecturers. In addition, an analysis of the dialogues and history of interaction between students and lecturers was performed. Finally, machine learning techniques were employed to identify patterns between student actions and lecturer diagnostics.

Porayska-Pomsta et al. (2008) distinguished 8 factors that were mainly employed by lecturers to provide feedback: student confidence, aptitude, effort, knowledge and interest; correctness of student answer; level of difficulty and importance of the material. Relevant affective states were distinguished: low confidence, effort and interest and confusion and happiness were reported, but in relation to student confidence. As a result, Porayska-Pomsta et al. (2008) focused on deriving decision trees that can infer student confidence, effort and interest using the relevant factors and their correlations. However, the sparsity of the data did not allow them to determine final rules that can be employed to provide feedback or diagnose student affective states.

3.3.15 Wayang Outpost with sensors in the classroom

Arroyo et al. (2009) focused on identifying student affective state with sensors, e.g. cameras and mouse, chair and wrist sensors, to identify physical features and physiological signals while teaching Mathematics in classroom settings. In addition, they investigated if student outcomes are related to affective states and if student emotion is related to their motivation and attitudes towards Mathematics. Arroyo et al. (2009) aim to identify relevant affective states in the educational context, e.g. joy, anger, surprise, fear, disgust, contempt and interest, based on Ekman's theory (Ekman 1999), which is related to analysing facial expressions.

The investigation was built over the Affective Agent Research Platform, previously created by Burleson (2006), which comprised a mouse, posture chair, camera and GSR bracelet in order to recognise and respond to affect using a learning companion. In the Affective Agent Research Platform, dialog and posture features were employed to recognise boredom, frustration, confusion and flow. In addition, Wayang Outpost (Arroyo and Woolf 2005) discussed in section 3.3.6 was modified through the incorporation of sensors and learning companions that focused on encouraging effort and perseverance and in parallel attempting to make students' aware of the available help and persuade them to make questions.

The investigation methodology by Arroyo et al. (2009) entailed enabling 38 students at high school level and 29 female students at undergraduate level to interact with Wayang Outpost in order to learn geometry. The student interaction with Wayang Outpost took four to five days. In parallel, students took a Mathematics test and survey investigating student motivation and perceptions for learning Mathematics. Results showed that both groups of participants improved their knowledge about geometry; about 10% on average after interacting with Wayang Outpost. In addition, the investigation showed that self-reported emotions were significantly related to the events, i.e. the context, which occurred while solving the preceding problem instead of being associated with student perceptions. As a result, contextual variables, such as the time elapsed and the gender of the learning companion, were employed to analyse if the reported emotions could be predicted significantly using these variables. Results corroborated this hypothesis, e.g. the number of hints asked in the preceding session accounted for 28% of the variance of frustration. Linear regression was employed to perform this analysis using the stepwise procedure.

Twenty five sets of sensors were employed to acquire student physiological and physical features, e.g. a chair with pressure sensitive cushions and accelerometers located in the seat and back in order to measure student position, GSR bracelet, facial recognition sensor with software to read mental states, such as concentration, and a pressure mouse. However, not all sensors were available for all the students during the study due to what the authors call

“real-life practical problems” (Arroyo et al. 2009, p. 22). As a result, only 50% of the data was obtained and employed for the analysis with linear regression, since if incomplete data is employed this approach would not work appropriately. Results showed that facial gestures contributed with approximately 60% of the variance of affective states that enhanced the accuracy of prediction when compared with using only contextual variables.

3.4 Demonstrating affect: emotional agents and robots

The new generation of ITSs also aim to demonstrate affect and to be believable, i.e. demonstrate a more natural and realistic behaviour (Johnson et al. 2000). The ultimate goal is to enhance the effectiveness of conveying an instructional message through adapting to student preferences and needs and managing student attention using AI techniques. Hence, they usually employ virtual peers or Embodied Pedagogical Agents (EPAs) that may be tutors or learning companions and NPCs with this objective as can be observed in the work of D’Mello (2008b), Arroyo et al. (2009) and McQuiggan et al. (2008) respectively. NPCs are virtual characters usually employed in GBL environments in order to enhance the delivery of the narrative and signal game goals (Collins 2008).

EPAs can have two roles: tutors (D'Mello et al. 2008b, Sarrafzadeh et al. 2008) or learning companions (Arroyo et al. 2009). In the latter the EPA is designed to be perceived as another student, i.e. with the same level of knowledge as the student. Hence, the student is expected to work as a team member with the learning companion, e.g. progressing together in knowledge and questioning each other’s knowledge and understanding. On the other hand when the EPA is designed as a tutor, it is expected that the EPA has a higher level of knowledge about the topic than the student. Therefore, the student is expected to work more independently and only request help and assistance when needed. As a result, EPA knowledge is considered as a believability feature (Johnson et al. 2000).

In the new generation of ITSs, in order to achieve believability, EPAs research have covered areas such as Emotional Intelligence, Common Sense, Sociology, Animation, Distributed Architectures, Multimodal Output Adaptation and Cinematography. Therefore, EPAs have been implemented using several goals and strategies, e.g. showing empathy and affinity (Sarrafzadeh et al. 2008), mirroring behaviour (Arroyo et al. 2009) or being polite (Porayska-Pomsta et al. 2008), since it is still not clear if the emotions that have a negative connotation, such as anger, actually have a negative influence on performance or learning. In fact, it is still not clear what the effects of emotion are and how negative emotions or affective states can be changed.

Also, research in affective educational modelling has sought solutions in the field of embodied conversational agents (ECAs), i.e. synthetic autonomous characters, and robotics.

The motivation behind ECAs is to observe that people employ verbal and nonverbal clues to express themselves and to communicate. Poggi and Pelachaud (2000) created a facial action coding system using the theory by Ekman and Friesen (1978) that associates specific movements of facial muscles with specific intentions in order to enable artificial agents to convey messages with facial expressions. The work by Dias et al. (2006) that focused on creating Fear Not!, which includes ECAs that perform bullying situations in order to assist children aged 8 to 12 on achieving an enhanced understanding of social problems is an example. Becker-Asano and Wachsmuth (2010) demonstrate how Damasio's (1994) primary and secondary emotions can be expressed by a virtual human, MAX, using a cognitive modelling approach, i.e. Belief-Desire-Intention (BDI). Primary emotions or innate emotions serve the purpose of giving a fast and reactive response of behaviour to dangerous situations. Becker-Asano and Wachsmuth (2010) decided to focus on conveying the nine primary emotions: anger, annoyance, boredom, concentration, depression, fear, happiness, sadness and surprise. Secondary emotions, e.g. relief or hope, arise from higher cognitive processes through the evaluation of preferences over outcomes and expectations. Furthermore, secondary emotions may influence body language as do primary emotions. Becker-Asano and Wachsmuth (2010) focused on the secondary emotions: hope, fears confirmed and relief. MAX, implemented as a multi-agent system, expresses as appropriate affective states to players in a card game scenario. Max's goal is to assist the player to follow the rules of the game and react to the game context accordingly.

Another example is the Sensitive Agent (SEMAINE), which is a multimodal system, a Sensitive Artificial Listener (SAL) capable of maintaining a conversation with human users (DFKI et al. 2011). It operates in real-time and employs facial expressions, gaze and voice to identify users' disposition or affective states. The user is engaged through a Embodied Conversational Agent (ECA) which considers the features in the users voice to adapt its response using its body language, facial expressions and voice modulation (see Figure 3.17). The agent also employs diverse strategies for engaging the user depending on the user's affective state.

Robotics is also endeavouring to incorporate physical intelligent devices, which interact with the world as people do, providing them with capabilities such as independent thinking, planning, natural language and sociability (Smith and Breazeal 2007). The rationale behind this is realising that current home robots, which serve mainly an utility purpose, are capable of arousing social responses from their owners (Breazeal and Brooks 2005). Cog and Kismet are two emotional humanoid robots from the MIT that learn collaboratively through social interaction with people (Turkle et al. 2006). Cog is a humanoid robot composed of visual, tactile and kinaesthetic sensors, whilst Kismet is a robotic head, which has facial features that can express basic emotions and can understand and generate speech. iCat is an emotional

and educational robot, the affective state of which is influenced by the player's moves on an electronic chess board, see Figure 3.18. It can predict the final state of the interaction. iCat shows emotional states that are related to the identification of good or bad moves, which aim to achieve an enhancement of the player's knowledge and understanding for playing chess (Pereira et al. 2008).

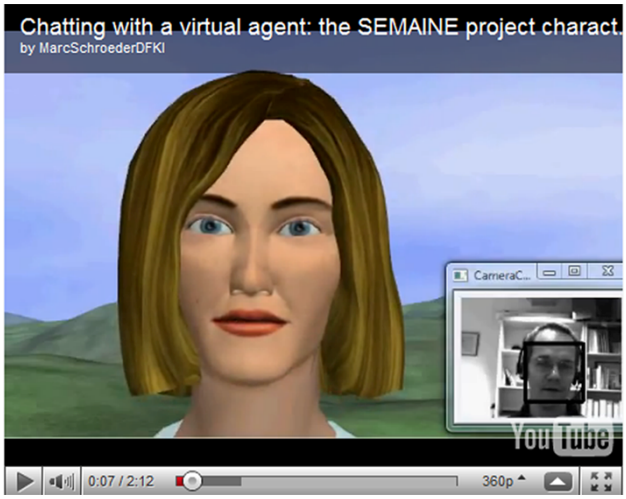


Figure 3.17 Interaction between a SAL and a human (DFKI et al. 2011)

Figure 3.18 iCat interacting with a player (Leite et al. 2008, p. 1231)

3.5 AI and machine learning techniques

Diverse methods are employed by this new generation of ITSs in order to achieve knowledge representation, a sub-area of Artificial Intelligence (AI), which enables computers to reason about concepts, e.g. student behaviour or the specific subject domain, and show intelligent behaviour (Woolf 2009). Rules are an example of these representations. The appropriate type of representation depends on the nature of the domain.

Rule-based systems are a technique for representing knowledge that, although common, are time consuming to implement and tend not to scale well. As example Anderson and Skwarecki (1986) had to define hundreds of rules that characterised every probable answer to student needs in LISP tutor. According to the specific problem, the system decides the rule that should be applied. Also, De Vicente and Pain (2002) derived 61 inference rules that correspond to every response that teachers will take in order to address student misconceptions and keep students motivated, while Jaques and Vicari (2007) use inference rules to create an emotional student model with the Ortony, Clore and Collins (OCC) model for diagnosing students emotion. Decision trees can assist in selecting the action that maximises the likelihood of achieving the predicted value of the output given a specific input through considering a set of attributes that define a given case or situation. Vectors and matrices are common data structures used for knowledge representation in the areas of robotics and synthethic characters to model affective and personality features (Miwa et al. 2001, Martinho and Paiva 2006). Miwa et al. (2001) implemented the WE-3RV humanoid robot, which senses a person's personality and attains effective communication using the Big Five Theory. A 3D mental space was defined where seven emotions (e.g. happiness, anger, disgust, fear, sadness, neutral and surprise) were mapped. When a stimulus is sensed by the humanoid-robot, the stimulus is processed to determine emotion. Then the emotion is expressed by selecting a personality from a matrix that defines six personalities based on the five factors of personality and assigning intensity variables (e.g. high and low) to each factor. The 3D mental model of WE-3RV is comprised of three axes: pleasantness, activation and certainty. The mental vector (M) represents the robot's own "mental state". Visual, auditory, tactile, temperature and olfactory sensations are processed by the mental model and the personality model. Martinho and Paiva (2006) used an "emotivector" as a mechanism of anticipation to enhance the believability of synthetic characters. An emotivector is comprised of a sensor that uses a history of states to determine the incoming state. They designed and implemented an agent architecture, which is used to manage multiple emotivectors. Each emotivector is related to one dimension of perception. The emotivectors are stored and managed by the Salience module. Every time new information is sent to the processing module, the related emotivector uses the history to compute the incoming value using a hybrid algorithm. The salience result generated by the emotivector generates an emotion. Finally, the new information is sent to the processing module to provide a recommendation. The emotivector concept was tested through the implementation of a puzzle game where the synthetic flower, Aini, monitors the players' progress and assists them to solve the word puzzle by reacting to their actions.

Plan recognition is employed in MORE (Del Soldato 1993) to recognise student goals and strategy and accordingly provide suitable feedback. To apply this technique interaction has to be designed at the level of actions. To facilitate the identification of patterns employed for

knowledge representation, data mining techniques can be employed in order to reduce uncertainty. According to Gartner (2011) data mining is defined as a process in which key correlations, patterns or trends are identified from large quantities of data in repositories. To facilitate the identification of knowledge from data, the Knowledge Discovery in Databases (KDD) process may be followed, which comprises seven steps (Han and Kamber 2006, Fayyad et al. 1996): (1) data cleaning, (2) data integration, (3) data selection, (4) data transformation, (5) data mining, (6) pattern evaluation and (7) knowledge presentation.

Data cleaning entails removing inconsistent data and noise. *Data integration* involves combining data located in various repositories. *Data selection* involves retrieving only relevant data from the repositories. *Data transformation* involves converting data through a series of aggregation or summary operations into more suitable forms for mining knowledge. *Data Mining* involves applying AI methods to extract patterns from data. Pattern evaluation involves assessing the extracted patterns according to *interestingness measures*, e.g. confidence threshold and intervals. Knowledge presentation involves presenting information to users through visualisation and representation techniques, e.g. IF-THEN rules or decision trees.

Machine learning methods are employed to identify patterns and relations in data and facilitate knowledge representation. Statistical methods may be employed for this purpose, but also to create a statistical model of data. Binary and Multinomial Logistic Regression (MLR) are some of the preferred statistical methods employed in psychology to classify category membership (Kinnear and Gray 2010). They are highly effective handling categorical regressors and very useful for modelling real world problems which cannot be reduced to orthogonal designs (Tabacknick and Fidell 2007). They are also helpful when correlations between independent variables (IVs) result from an unequal number of cases, i.e. they do not have to comply with multivariate normality nor homogeneity of variance-covariance matrices (Kinnear and Gray 2010). MLR facilitates evaluating the relationship between one dependent variable (DV) and a number of IVs, which are correlated with one another and with the DV in varying degrees. The advantages of Binary and MLR are enabling visualisation of the contribution of each IV to the odds in favour of having the DV. Next, we explain the process of logistic regression.

The *odds* in favour of an event are defined as ‘the number of ways in which an event can occur divided by the number of ways in which the event can fail to occur’ (Kinnear and Gray 2010, p. 551). For example if a student can report 1 out of 4 emotions (enjoyment, anger, frustration or boredom), the odds in favour of frustration are 1/3. The *probability* of an event is the number of ways in which the event can occur divided by the total number of possible outcomes. In this case the probability that the student reports frustration is 1/4. The odds and the probability of an event are related by the expression in Equation 3.1. The *logit* is the natu-

ral logarithm of the odds and it is employed because, in comparison with the odds, the logit has a symmetry of range, e.g. at 50/50 the odds are equal to 1 (events with odds greater than 1 are categorised as 'likely' and events with odds lower than 1 are categorised as 'unlikely'), whilst the logit of 1 is zero at 50/50.

$$p = \frac{odds}{1 + odds} \qquad \text{Eq. 3.1}$$

In logistic regression, it is assumed that the logit is a linear function of the independent variables (see Equation 3.2). Therefore, the logistic regression equation can be defined as shown in Equation 3.3. The values of the parameters b_0, b_1, b_2, ..., b_n – *logistic regression coefficients b* - are chosen to enable the logistic equation to classify the independent variable, i.e. category membership, as precisely as possible. There is no mathematical solution to the challenge of determining the logistic regression coefficients. As a result an algorithm, based on a series of iterations is employed to converge upon stable estimates. Each logistic regression coefficient represents the increase in the logit by an increase of one unit in the independent variable, i.e. the log of the odds increases by *b* units. In the Statistical Package for the Social Sciences (SPSS) (IBM 2012), there are three methods to include independent variables in the logistic regression equation: (1) the simultaneous method, (2) the hierarchical method and (3) the stepwise procedure or forward conditional. In the first case, all the variables are introduced at the same time. In the second case, control variables are introduced in the analysis before the regressors whose effects are of primary concern. In the third case, the independent variables are selected in the order in which they maximise the statistically significant contribution to the model.

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = b_0 + b_1 X_1 + b_2 X_2 + ... + b_n X_n \qquad \text{Eq. 3.2}$$

$$\hat{p} = \frac{e^{\ln(odds)}}{1 + e^{\ln(odds)}} = \frac{e^{b_0 + b_1 X_1 + b_2 X_2 + ... + b_n X_n}}{1 + e^{b_0 + b_1 X_1 + b_2 X_2 + ... + b_n X_n}} \qquad \text{Eq. 3.3}$$

Other examples of knowledge representations are Artificial Neural Networks (ANNs), Bayesian Belief Networks (BBNs) and Dynamic Bayesian Networks (DBNs). DBNs have been employed in the computer instruction field to reason over time about student misconceptions (Goguadze et al. 2011) and student affective states (Conati and Maclaren 2009, Sabourin et al. 2011). Probabilistic models also have been successfully employed in the robotics field (Thrun 2002). ANNs can find patterns between the emotion inferred by expert

judges in videos and physiological signals in AutoTutor (D'Mello et al. 2008b). BBNs and DBNs are highly effective for handling uncertainty and simultaneously are employed to characterise prior domain knowledge and domain dependencies (Jensen and Nielsen 2007). DBNs assist in achieving information about the likelihood of students knowing a topic or experiencing an emotion before or after the interaction (Conati and Maclaren 2009, Sabourin et al. 2011). In addition, they are a form of knowledge representation that are effective for handling the uncertainty of a domain that evolves over time. *Influence diagrams* can be employed to enhance BBNs through including utilities and action nodes (Russell and Norving 2003), which assist in determining the present state, and include awareness of probable actions and their potential outcomes and the utility of each state.

DBNs, in the traditional sense, are directed graphical models of stochastic processes. They generalise Hidden Markov Models (HMMs) and Linear Dynamic Systems (LDSs), since they represent the hidden state in terms of state variables that may have complex interdependencies. A more appropriate term for DBN would be *temporal Bayesian network*, since it is assumed that a traditional DBN model is time-invariant, i.e. its structure and parameters do not change, but the term DBN has persisted (Murphy 1998). On the other hand, with complex DBNs, extra nodes may be added to represent the current state of the system more appropriately, which creates a combination of models in order to capture periodic non-stationaries. Sometimes even the size of the space can change over time, which may involve changing the structure of the model over time. Figure 3.19 shows a complex DBN, a model of an automated taxi, and demonstrates how *slice* t differs from *slice* $t+1$ in its structure, i.e. the number of random variables and relationships has increased in order to represent the system more appropriately.

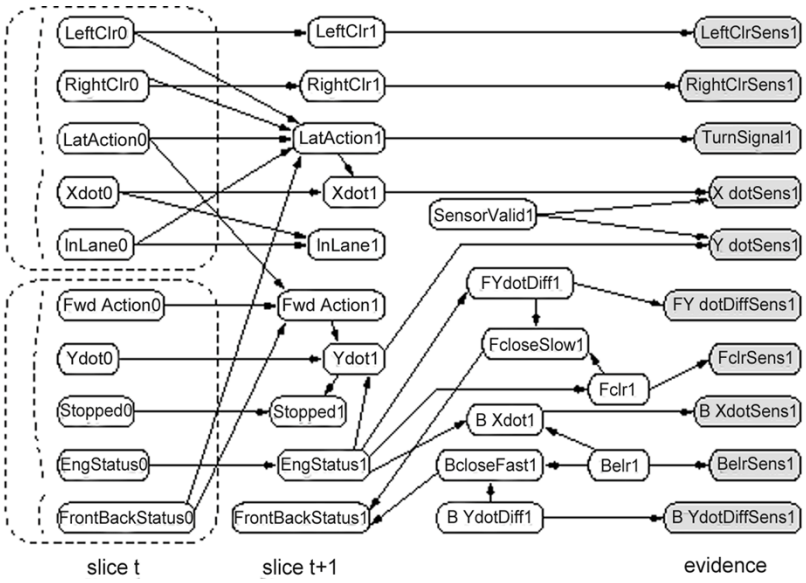


Figure 3.19 Bayesian Automated Taxi Network (Murphy 1998)

The DBNs approach originated due to the fact that real world events comprise temporal relationships between different entities, which may be described through multiple states of observation, i.e. time-series analysis or stochastic processes (Mihajlovic and Petkovic 2001). BBNs do not provide a direct mechanism for representing temporal dependencies. For including or adapting the temporal dimension in BBNs several approaches have been proposed. This temporal dimension is referred to in BBNs as 'temporal' or 'dynamic'. *Temporal models* focus on modelling time as a continuous permanent category, while *dynamic models* focus on modelling other changes in the system, such as a change in state.

There are several types of DBNs, such as *pure probabilistic DBNs*, which are based only on a probabilistic framework, i.e. we have just states, as objects and arcs representing conditional dependence between states from the same time slice or from two consecutive time slices (Mihajlovic and Petkovic 2001). These temporal arcs have the same nature as the arcs in static BBNs, but they represent conditional dependencies that exist between states in different time instances. As a result, DBNs are usually treated as a special case of BBNs. In a DBN, the variables denote its state, since a DBN represents the temporal dimension. The states of a DBN satisfy the Markovian condition or first order Markov property that may be stated as follows: The state of a system at time t depends only on its state at time $t\text{-}1$ (the future is independent of the past given the present), i.e. the transition model for first-order Markov processes ($\Pr(x_t|x_{t-1})$).

A DBN comprises a probability distribution function (see Equation 3.4) over the sequence of T state variables $X = \{X_0, \ldots X_{T-1}\}$ and the sequence of T observable variables $Y = \{Y_0, \ldots, Y_{T--1}\}$, where T is the time boundary for the given event under research. Therefore, to completely specify a DBN, it is necessary to define three sets of parameters: (1) time dependencies between states, i.e. state transtitions probability density functions $\Pr(x_t|x_{t-1})$, (2) observation nodes dependencies concerning other nodes at time t, i.e. observation probability density functions $\Pr(y_t|x_t)$ (sensor model) and (3) the initial probability distribution at the beginning of the process, i.e. the initial state distribution $\Pr(x_0)$. The first two parameters had to be determined for all states in all time slices $t = 1, \ldots, T$.

$$\Pr(X,Y) = \prod_{t=1}^{T-1} \Pr(x_t \mid x_{t-1}) \prod_{t=0}^{T-1} \Pr(y_t \mid x_t) \Pr(x_0) \qquad \text{Eq. 3.4}$$

An approach for restricting the parents of the state variables (X_t) involves restricting the parents of the evidence variables (Y_t) (Russell and Norvig 2010). It is assumed that the evidence variables at slice t depend only on the current state. Equation 3.5 illustrates this approach, where the conditional distribution $\Pr(Y_t|X_t)$ is known as the observational model or sensor model and it describes the manner in which the evidence variables or sensors are in-

fluenced by the current state of the world. The direction of the arc is always from state to sensor.

$$\Pr(Y_t \mid X_{0:t}, Y_{0:t-1}) = \Pr(Y_t \mid X_t) \qquad \text{Eq. 3.5}$$

Conditional probability density functions can be time invariant ($\Pr(x_t|x_{t-1}) = \Pr(x_t|x_{t-1}, t)$ or variant and time invariant conditional probability density functions can be parametric, ($\Pr(x_t \mid x_{t-1}) = \Pr(x_t \mid x_{t-1}, \theta)$ or non-parametric when they are represented using Condional Probability Tables (CPTs). Also depending on the type of state space of variables that needs to be represented, DBNs may be continuous, discrete or a combination of the two.

A type of *pure probabilistic DBN* is *an extension of a BBN towards a DBN*. This enhancement may be attained in several ways, which are classified in five categories (Mihajlovic and Petkovic 2001, Singhal and Brown 1997):

1. The addition of a history node to the corresponding node in the BBN, e.g. $\Pr(X_t \mid X_{t-1})$, see Figure 3.20

2. Run time selection of BBNs that involves the pre-development of a library comprising fully structured DBNs, which are selected using as a basis a specific state of the System

3. Dynamic structural changes in the BBN, which more accurately model human reasoning and are triggered when beliefs in one node or nodes in the DBN structure exceed a specific threshold value. In Figure 3.21, the Road node is added to the car and road detection BBN using vision and noise sensors when the probability corresponding to the car presence reaches a threshold larger than 0.5

4. Uniform representation of duplicated BBNs for every time slice, with introduction of BBNs for events representation. The nodes that correspond to these event BBNs receive evidence from nodes of current and previous time slices

5. Uniform representation of duplicated BBNs, which include additional temporal arcs among nodes that depend on some other nodes from previous time slices. Since these connections may lead to highly complex structures, in the majority of the cases only the connections from the previous time slice to the current time slice are allowed.

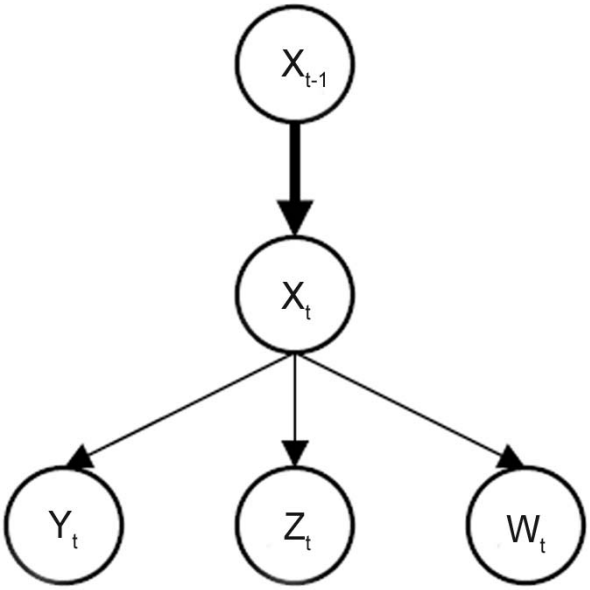


Figure 3.20 Addition of a history node in a BBN (Singhal and Brown 1997, p.7)

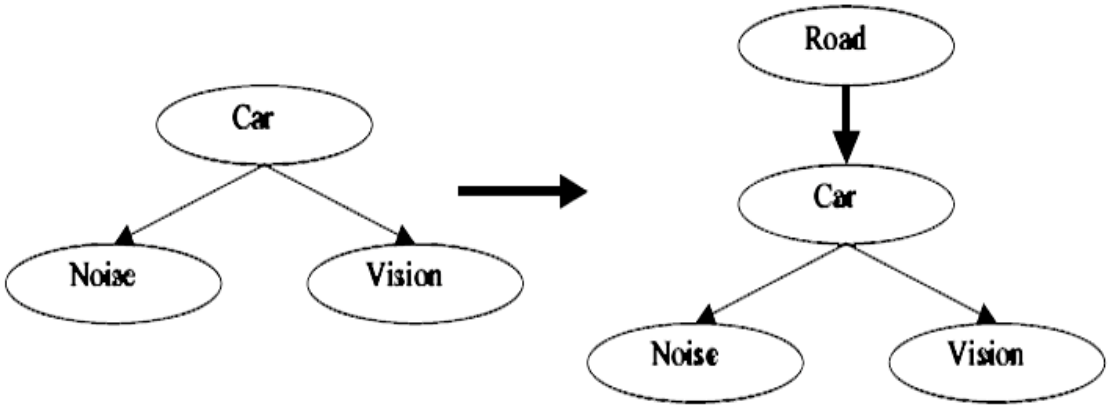


Figure 3.21 Dynamic structural changes in a BBN (Singhal and Brown 1997, p.7)

The first three methods of extension for BBNs have been applied by Singhal and Brown (1997) for reasoning about sensor discrepancies in robotic navigation in generalised environments. DBNs that correspond to the first category can be considered as BBNs with additional nodes. Run time selection BBNs, the second category of BBNs, involves selecting a subset of BBNs that match the beliefs about the entities in the world or environment. The system is static in the sense that it cannot learn new DBNs and operates with a known set, but the system is also dynamic, since it can change its own structure and utilise another DBN in the following time instance. The third category corresponding to structural changes in DBNs can take place on CPT values over time or including and removing arcs or nodes. Decomposition techniques of real world situations can be applied to achieve this goal, i.e. sometimes it is more convenient to create submodels for each stage (a complex temporal

model is composed of several simpler models). Through composition the number of models is reduced. Each submodel is learned through observations in specific time instances.

DBNs for representing events, corresponding to the fourth category of extension, use obtained information from different sources to infer events that are happening between two consecutive time points or slices in order to determine the action that is happening among both states (Mihajlovic and Petkovic 2001). Figure 3.22 shows this kind of network, which comprises three types of nodes: world nodes (W), event nodes (E) and observation nodes (O). World nodes represent the central domain variables. Observation nodes represent direct observations of the world and event nodes represent a change in the state of the world. On this kind of model, time intervals among time slices are not unique and depend on the occurrence of discrete events. Each time slice represents a static environment during that time interval, because the changes do not occur very often and the structure within that time slice is frequently regular. Two types of observations are present in DBNs, see Figure 3.22: (1) Direct observations of a world variable $O(T)$ and (2) Observations of an event $O(T_i, T_{i+1})$. This model has been successfully applied on fall monitoring walking and fall diagnosis. The fifth category of extension, DBNs with a uniform structure, (see Figure 3.23), comprise an identical structure for every time slice and identical temporal conditional dependencies (arcs) between time slices, i.e. conditional dependencies and states in each time slice are a BBN and the DBN is constructed over these BBNs, by multiplying them for each time slice and by incorporating arcs among states of consecutive time slices if they are temporally correlated.

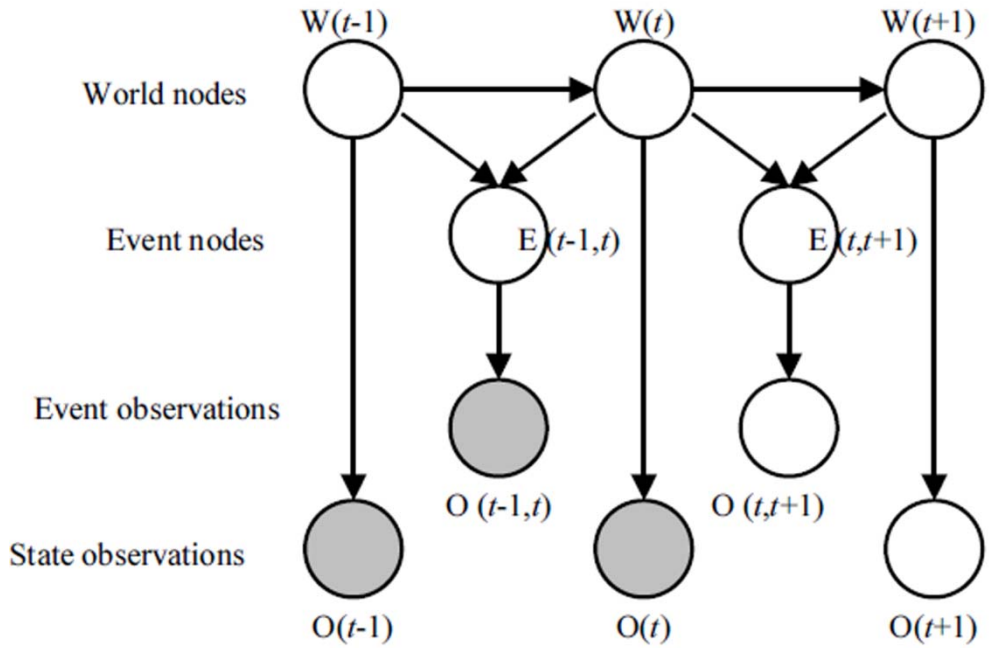


Figure 3.22 DBNs for representing events (Mihajlovic and Petkovic 2001, p.26)

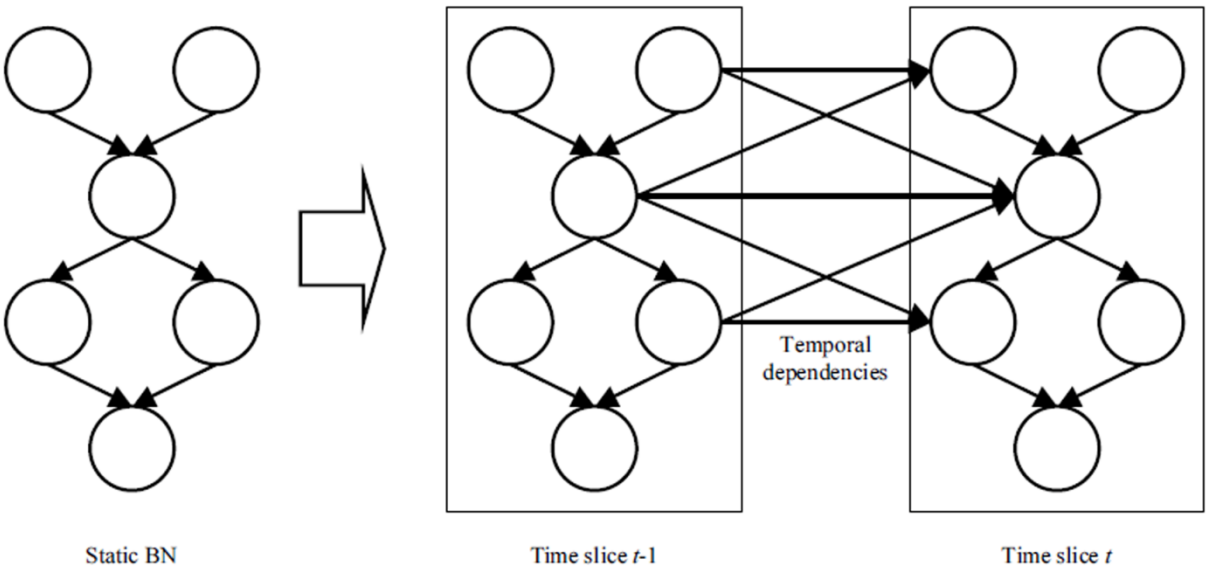


Figure 3.23 DBN with a uniform structure (Mihajlovic and Petkovic 2001, p.26)

The approach followed in our research to define a dynamic sequence of BBNs is similar to the second enhancement of BBNs towards DBNs described by Mihajlovic and Petkovic (2001) where a BBN is selected from a library by considering the state of the System. However, it also shares similarities with a DBN approach, since BBNs comprise historic nodes and relations of the form $\Pr(X_t \mid X_{t-1})$.

Binary/MLR and Bayesian models have been employed for discriminative prediction, e.g. classification. Roos et al. (2005) show that there is a relationship between the parameter learning methods of Bayesian models and logistic regression. As a result, any conditional Bayesian Network model satisfying specific conditions may be equivalent to a logistic regression model. An explanation of this statement follows: First, it is important to indicate that Roos et al. (2005) provide a new parameterization of Bayesian network models ($\mathcal{M}^{\mathcal{B}}$), so the parameters correspond to logarithms of parameters in the standard Bayesian network parameterization and as a result, each conditional Bayesian network model can be mapped to a logistic regression model. Let X_0 be a random variable with positive values $\{1, ..., n_0\}$, and let $\mathbf{Y}= (Y_1, ..., Y_k)$ be a real-valued random vector. The logistic regression model with independent variable X_0 and covariates $Y_1, ..., Y_k$ is defined as the set of conditional distributions (see Equation 3.6) where the parameter vector β with components of the form $\beta_{X0,i}$ with $X_0 \in \{1, ..., n_0\}$, $i \in \{1, ..., k\}$ is allowed to take on all values $\Re^{no\ k}$. $\Re^{no\ k}$ is a convex set in which the log-likelihood is allowed to vary freely. This is owed to employing the new parameterization defined by Roos et al. (2005). As a result, the log-likelihood becomes a concave function of the parameters under the perfectness condition. The latter implies that the global maximum can be found in the conditional likelihood surface by simple local optimisation techniques,

e.g. hill climbing, but also there is the possibility that there are not network structures for which the conditional likelihood surface has local, non-global maxima.

$$P(X_0 \mid y, \beta) := \frac{\exp \sum_{i=1}^{k} \beta x_i, iY_i}{\sum_{x'_{=o}=1}^{n_0} \exp \sum_{i=1}^{k} \beta_{x'_{=o}}, iYi} \qquad \text{Eq. 3.6}$$

For all the values of the class variable $r \in \{1, ..., n_0\}$ and all covariates $s \in \{1, ..., k\}$, the partial derivatives of the log-likelihoods, i.e. components of the gradient vector are given by Equation 3.7, where $\boldsymbol{I}_{[r=xo]}$ is the indicator function (if the argument is true takes a value of 1, 0 otherwise). The Hessian matrix (H) of second derivatives has entries given by the expresion in Equation 3.8, where $r,t \in \{1, ..., n_0\}$ and $s,u \in \{1, ..., k\}$. By applying the theorem that states that the Hessian Matrix is negative semidefinite (McLachlan 1992), the model is extended to several outcomes under the independent and identical distributed (i.d.d.) assumption by defining the log likelihood as shown in Equation 3.9.

$$\frac{\partial \ln P(X_0 \mid y, \beta)}{\partial \beta_{r,s}} = Y_S \left(I_{[r=X_0]} - P(r \mid y, \beta) \right) \qquad \text{Eq. 3.7}$$

$$\frac{\partial^2 \ln P(X_0 \mid y, \beta)}{\partial \beta_{r,s} \partial \beta_{t,u}} = -Y_S Y_u P(r \mid y, \beta) \left(I_{[r=t]} - P(t \mid y, \beta) \right) \qquad \text{Eq. 3.8}$$

$$CLL_L(D; B) := \sum_{j=1}^{N} \ln P\left(x_0{}^j \mid y^j, \beta \right) \qquad \text{Eq. 3.9}$$

Given a data set, the entries of the gradient vector and the Hessian Matrix are sums of terms given by Equations 3.7 to 3.8 respectively. But, by applying the previously introduced theorem related to the negative semidefinitive character of the Hessian Matrix for each data vector, the logarithm of Equation 3.6 is concave and the log-likehood of a sum of concave functions is also concave, i.e. a concave function of the parameters, and when the fact that the parameter vector β varies freely in a convex set ($\Re^{no\ k}$) is considered, it is observed that there are no local or global maxima in the log-likelihood surface of a logistic regression model.

Taking into account the previous analysis of logistic regression models, Ross et al. (2005) showed how to create a logistic model that corresponds to a Bayesian Network. Consider that we have a Bayesian Network in which it is assumed that all the parents of the class variable are fully connected (see Figure 3.24 (a)). The BBNs in Figures 3.24 (a) and 3.24 (b) are

equivalent in terms of conditional distributions for the random variable *disease* given the shaded variables. Figure 3.24 (a) is the canonical structure ($\mathcal{M}^{\mathcal{B}^*}$) of the Bayesian Network ($\mathcal{B}^*$), while Figure 3.24 (b) is the arbitrary structure ($\mathcal{M}^{\mathcal{B}}$) of the Bayesian Network ($\mathcal{B}$). The former is obtained when the latter is restricted by X_0's Markov Blanket and arcs are added as necessary to make all the parents of X_o fully connected.

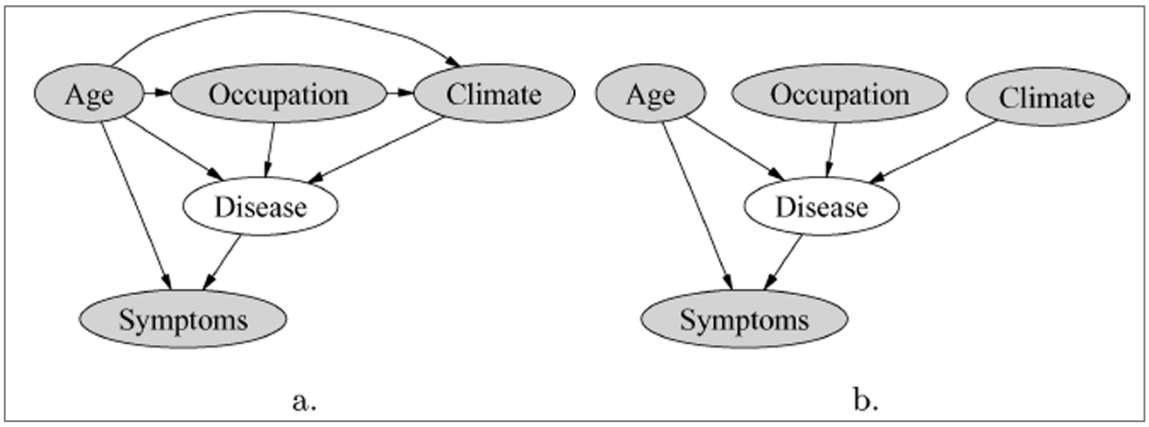


Figure 3.24 Equivalent Networks (Roos et al. 2005, p. 274)

Ross et al. (2005) defined sets of covariates shown in Equation 3.10 and 3.14 for all parent configurations $\boldsymbol{pa}_0$ of X_0. The parameters associated with the covariates shown in Equation 3.10 are denoted as shown in Equation 3.11. Equation 3.12 defines the parent set of X_i with the exclusion of the class X_0, where $i \in \{1, ..., m\}$, $x_i \in \{1, ..., n_i\}$ and considering the definition in Equation 3.13. The parameters associated with the covariate in equation 3.14 are denoted as shown in Equation 3.15.

$$Ypa_0 := I[Pa_0 = pa_0] \qquad \text{Eq. 3.10}$$

$$\beta^{\mathcal{B}}_{X_0, pa_i} \qquad \text{Eq. 3.11}$$

$$Pa_i^+ := Pa_i \setminus \langle X_0 \rangle \qquad \text{Eq. 3.12}$$

$$pa_i^+ \in dom(Pa_i^+) \qquad \text{Eq. 3.13}$$

$$Y_{X_I}, pa_i^+ := I_{[X_i = x_i, Pa_i^+ = pa_i^+]} \qquad \text{Eq. 3.14}$$

$$\beta^{\mathcal{B}}_{x_0, x_i pa_i^+} \qquad \text{Eq. 3.15}$$

Considering the defined sets of covariates and the BBN in Figure 3.24 (a), the covariates of the type described by Equation 3.10 correspond to all the combinations of values related to the random variables *age, occupation* and *climate* and the covariates of the type described by Equation 3.14 correspond to all the combinations of values related to the random vari-

ables *age* and *symptoms*. The logistic model obtained is shown in Equation 3.16, following the notation of Equation 3.6, this model was written as a function of variables $X_1, ..., X_M$. The same model would be obtained from Figure 3.24 (b). Equation 3.17 is the simplified version of Equation 3.16, since most of the indicator variables take a value of zero.

$$P\left(x_0 \mid x_1,...,x_M,\beta^{\mathcal{B}}\right) := \frac{\exp\left(\sum pa_0 \beta^{\mathcal{B}}_{x_0,pa_0} y_{pa_0} + \sum_{i=1}^{m}\sum_{x_i=1}^{n_i}\sum_{pa_i^+} \beta^{\mathcal{B}}_{x_0,x_i,pa_I} y_{x_i,pa_i^+}\right)}{\sum_{x_0'=1}^{n_0}\exp\left(\sum_{pa_0} \beta^{\mathcal{B}}_{x_0',pa_0} y_{pa_0} + \sum_{i=1}^{m}\sum_{x_i=1}^{n_i}\sum_{pa_i^+} \beta^{\mathcal{B}}_{x_0',x_1,pa_i^+} y_{x_i,pa_i^+}\right)} \quad \text{Eq. 3.16}$$

$$P\left(x_0 \mid x_1,...,x_M,\beta^{\mathcal{B}}\right) := \frac{\exp\left(\beta^{\mathcal{B}}_{x_0,pa_0(x)} + \sum_{i=1}^{m}\beta^{\mathcal{B}}_{x_0,x_i,pa_I(x)}\right)}{\sum_{x_0'=1}^{n_0}\exp\left(\beta^{\mathcal{B}}_{x_0',pa_0(x)} + \sum_{i=1}^{m}\beta^{\mathcal{B}}_{x_0',x_1,pa_i^+(x)}\right)} \quad \text{Eq. 3.17}$$

Let the conditional model $\mathcal{M}_L^{\mathcal{B}}$ be the set of conditional distributions for X_0, which can be represented by the logistic regression model corresponding to $\mathcal{B}$, the model $\mathcal{M}_L^{\mathcal{B}}$ Is very closely related to the corresponding BBN model $\mathcal{M}^{\mathcal{B}}$ as described by Theorem 2: 'Let $\mathcal{M}^{\mathcal{B}}$ be the set of conditional distributions that can be represented by a Bayesian Model with network structure $\mathcal{B}$ and strictly positive parameters, and let $\mathcal{M}_L^{\mathcal{B}}$ be the conditional model defined by the logistic regression model with covariates in Equations 3.10 and 3.14 . Then $\mathcal{M}^{\mathcal{B}} \subseteq \mathcal{M}_L^{\mathcal{B}}$' (Roos et al. 2005, p. 277).

Inference can be defined as finding the most likely explanation for a set of observations (Glass 2012). Two types of reasoning or inference are available in Bayesian models: (1) predictive or deductive reasoning and (2) diagnostic or abductive reasoning. The former is when the target variable is always a descendant of the evidence and the latter is when the target variable is usually an ancestor of the evidence (Jensen and Nielsen 2007). The propagation of evidence, knowing the probability for each state of a given variable from the joint probability distribution over all the variables but maxing out the other variables, is achieved by employing a propagation algorithm such as max-propagation (Hugin Expert A/S 2012).

Figure 3.25 summarises the methodology employed to define Bayesian models. For defining its structure, i.e. skeleton, it is necessary to know the conditional independent or dependent relations (CIDRs), which can be defined with assistance of a domain expert or obtained through statistical tests of historical domain data using a learning algorithm such as Peter-Clark (PC) or Necessary Path Condition, which depends on the amount of data that can be acquired and that will be employed to derive, train and evaluate the model.

PC algorithm statistically tests for conditional independence and derives a network skeleton. This goal is achieved by applying the principle of Occam's Razor: selecting the simplest model among equally good models (Jensen and Nielsen 2007). The PC algorithm is particularly effective on large data sets. The Necessary-Path Condition algorithm is suitable when data sets are limited, since various independence relationships may be found for a pair of variables and the user can participate in deciding the directionality of relationships, as a result solving uncertain relations. Once the structure of a Bayesian model is defined, it then proceeds to achieve parameter learning from data. Probabilities of the Conditional Probability Tables (CPTs) are learned in case of discrete chance nodes of Bayesian models, and in the case of continuous chance nodes probability densities are determined. However, the majority of algorithms employed for parameter learning are focused on discrete chance nodes. Hence, continuous random variables are usually exposed to a discretisation procedure in order to transform them into categorical random variables.

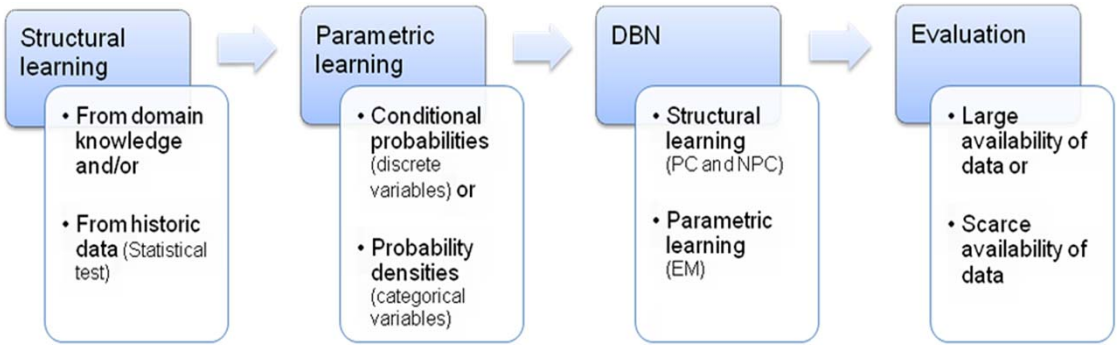


Figure 3.25 Methodology employed for deriving Bayesian models

The PC and Necessary Path Condition algorithms are based on a conditional independence test and are available for structural learning in a tool for learning BBNs, also known as Hugin Lite (Madsen et al. 2003). However, there are more methods available in a free-ware data mining tool, known as the Waikato Environment for Knowledge Analysis (WEKA) (The University of Waikato 2012). These methods are mainly classified in the areas of local score metrics, conditional independence tests and global score metrics (Bouckaert 2008). The local score metrics approach involves learning a network structure given the training data $Q(B_S|D)$, which has to be maximised. The quality measure may be based on a Bayesian or other approach, such as minimum description length information. The score of the individual nodes (as a sum or product) comprises the score of the whole network. The conditional independence test approach involves identifying a causal structure. As a result, it is assumed that there is a network structure that precisely represents the data distribution, i.e. if there is a conditional independency between two variables, there should not be any association between these two variables. The global score metrics approach involves estimating expected

utilities, such as calculating the average performance of the BBN (a score) over the testing sets. The whole network has to be considered to determine this metric. Diverse algorithms have been implemented in WEKA corresponding to these areas, e.g. hill climbing, simulated annealing and tabu search. The Expectation Maximisation (EM) learning algorithm learns the probabilities for discrete chance nodes from observed data, which does not necessarily require completeness (Jensen and Nielsen 2007).

A method employed for facilitating knowledge representation, specifically the derivation of BBNs and DBNs as a means of characterising a student model, is the application of Probabilistic Relational Models (PRMs). PRMs are object representations of the domain (Sucar and Noguez 2008), i.e. the domain can be characterised as series of entities with properties and relationships between them. This approach is based on the ideas of Koller (1999) where objects are grouped into disjoint classes, $X_1,...,X_n$, with dependencies determined at class level. Each class comprises specific attributes $A(X_i)$ and attributes can take precise values, $A(X_i) = A_j$, of a predetermined domain of values $V(A_j)$. Noguez and Sucar (2005) employed PRMs to characterise their student model comprising knowledge about mobile robotics that the students should achieve, their characteristics, previous performance and contextual variables corresponding to students interaction with the VLE.

3.6 Affective model evaluation

The evaluation of affective models and systems has two goals: (1) ensuring that the emotion experienced by the student is accurately classified or predicted or (2) guaranteeing that the emotional or affective state expressed by the system was perceived as believable or genuine. There are two approaches to evaluating affective student models. An approach for assessing the classification accuracy of the affective student model, employed by Conati and Maclaren (2009), combines micro-average and macro-average techniques. Micro-average involves calculating the percentage of cases correctly classified over all the instances without taking into consideration the class. Macro-average entails calculating the percentage of cases correctly classified per class.

Another approach (Arroyo and Woolf 2005, D'Mello et al. 2008b, McQuiggan et al. 2008) is cross-validation, which is the preferred technique when data is limited in practical situations (Witten and Frank 2005). It involves holding specific sets of data for training and testing. For example, two thirds of the data is employed for training and the other third is employed for testing. However, to ensure that all the classes are unevenly represented in the training and testing data sets, random sampling must be performed on the data sample. This is known as the *stratification procedure*.

The evaluation method used depends on the quantity of data available. If there is a large quantity of data then a large subset can be employed for training and testing. But both subsets have to be mutually independent. This method is referred as *hold-out estimate* (Bouckaert et al. 2012). However, when data is scarce, *cross-validation* is employed, where the dataset is divided into *n* sub-samples and one of these is held for testing the model and the *n-1* sub-samples are employed for training. For example if three folds are decided, data is divided into three equal partitions. Whilst two stratified partitions are employed for training, one is employed for testing. The procedure is repeated, in this case three times, until every instance has been employed for testing. The error rate is calculated in the instance used for testing. Finally, the three error estimates are averaged to give the overall error. The standard is to use 10 folds, since research has shown that it is the most suitable value to determine the best error estimate. However, a single 10-fold cross-validation is sometimes not enough to achieve accurate error estimation. Sometimes this involves running the algorithm 100 times with data sets that are 90% the size of the original, which involves intensive computation and large amounts of data.

Embodied Pedagogical Agents (EPAs), synthetic characters or Non-Player Characters (NPCs) are employed in order to communicate an affective state. To understand if an emotional or affective display was conveyed to students without becoming biased or confused in the process it is important to understand their context and cultural differences and the common patterns of reaction experienced when interacting with this kind of application. The ultimate goal is to achieve an effective design and facilitate the acceptance of the application. Quantitative-scientific and open-ended interpretation methods are employed to understand student perceptions and evaluate these applications (Höök 2005). The former has problems achieving a detailed view of the experience of end-users, whilst the latter assists in achieving knowledge that is temporary and culturally dependent, i.e. user-specific, but that cannot be generalised to a specific population. The most common problems in the design and implementation of affective applications are in the synchronization, contextualisation, user interaction control, timing and realism. The latter depends on user expectations of the response that avatars that look like humans should be capable of offering.

3.7 Comparison of new generation ITS systems

Tables 3.1 and 3.2 summarise research approaches, machine learning techniques and AI tools employed and features of the new generation of ITSs applications. The new generation of ITSs has the ultimate goal of keeping and encouraging student engagement and interest in parallel to the initial goal that ITSs have had from their creation, which is achieving knowledge, learning and understanding effectively. From the student modelling viewpoint, this has

entailed investigating the creation of ITSs capable of identifying students' preferred learning styles (Kelly and Tangney 2002, Souto et al. 2002), diagnosing students' motivation (De Vicente and Pain 2002, Del Soldato and Du Boulay 1995, Rebolledo-Mendez et al. 2006), recognising students' attitudes (Arroyo and Woolf 2005), inferring students' levels of self-efficacy (McQuiggan et al. 2008) and more recently inferring students' affective or emotional states (Conati and Maclaren 2009, D'Mello et al. 2008b, Jaques and Vicari 2007, Schaller et al. 2005). The last case has received increased attention due to recent work that has shown that emotional or affective states influence student motivation, decisions and performance (Picard 1995, Picard et al. 2004). Also Schutz and Pekrun (2007) have stated that lecturers may be affected in the same manner.

3.7.1 Approaches to recognising emotion

New generation ITSs employ cognitive or situational theories with approaches to modelling student knowledge or understanding resulting in adaptive instruction. Table 3.2 shows that there are three main approaches to recognising or reasoning about an affective state: (1) identifying the physical or physiological effects of emotion, (2) reasoning about observable behaviour from its origin and (3) hybrid approaches combining both.

Identifying the physical or physiological effects of emotion entails using hardware equipment, e.g. cameras and body posture sensors, heart rate or pressure mouse sensors, in order to acquire features and information related to student facial gestures, eye movement, body language or physiological signals in order to find relationships or patterns that can classify or predict students' affective states (Burleson and Picard 2007, D'Mello et al. 2008b, Neji and Ben Ammar 2007, Sarrafzadeh et al. 2008). Usually the opinion of expert judges, e.g. lecturers, or student self-reporting is employed to map student features to emotions. Machine learning techniques are employed to find these patterns or relations. This approach has proven to be the most successful thus far. However, it is still not clear which sources of information should be taken into account or if all sources of information should be employed in parallel (D'Mello et al. 2008b). However, this approach has only been plausible in laboratory settings and not in the place where learning actually happens, e.g. classrooms or online, since not all students have access to the expensive hardware equipment required, which is also prone to failure as can be observed in the work by Arroyo et al. (2009). In addition, processing these features require high bandwidth that deteriorate the system's performance. In addition, facial gestures have shown to be the most successful and useful features for recognising emotion, since they have shown to provide a significant amount of information (Arroyo et al. 2009, Westerinck et al. 2008b) about valence and arousal. Ekman (1999) and

	Learning Environment Characteristics														
	System	Open Learning Environment		Educational-Artificial Intelligence		Teaching Subject	Educational Level	Knowledge Represented			Reasoning / Recognising Approach		AI Approach	Underlying Theoretical Approach
		GBL environment	VLE	ITS	Intelligent Agents			Student's Cognitive Psychology Features	Student's Physical/ Biofeedback Features	Contextual Variables	Identification of Physiological Effects	Reasoning from Origin		
SELF EFFICACY	CRYSTAL ISLAND Learning Environment (McQuiggan et al. 2008)	✓	✗	✓	✗	Microbiology & genetics	Undergraduate	✓	✓	✓	✓	✓	Probabilistic models & rule-based models	Observations of student interactions & the self-efficacy problem solving instrument (Bandura 2006)
								(related to self-efficacy)						
MOTIVATION	M-Ecolab (Rebolledo-Mendez et al. 2006)	✓	✓	✗	✓	Ecology (Food chains & webs)	Primary	✓	✗	✓	✗	✓	Production rules	Motivational strategies signalled by Keller (1983) and Malone & Lepper (1987)
								(related to confidence, independence & effort associated with motivation)						
	MOODS (De Vicente 2003)	✗	✓	✓	✗	Japanese numbers	Postgraduate	✓	✗	✓	✗	✓	Inference rules	Motivational strategies signalled by Keller (1983) , Malone & Lepper (1987) & suggestions of post-graduate students with/without teaching experience
								(related to independence, attitude to challenge, control, expertise, effort, confidence, sensory & cognitive interest, satisfaction & relevance to student's goals associated to motivation)						
	MORE (Del Soldato & Du Boulay 1995)	✗	✓	✓	✗	Computer Science (Prolog language)	Postgraduate	✓	✗	✓	✗	✓	Production rules	Motivational strategies signalled by Keller (1983) and Malone & Lepper (1987)
								(related to confidence, independence & effort associated with motivation)						
ATTITUDES	Wayang Outpost (Arroyo & Woolf 2005)	✗	✓	✓	✗	Mathematics	High school	✓	✗	✓	✗	✓	Bayesian Belief Network (BBN)	Data logs of student interaction with an ITS
								(related to attitude towards learning)						
LEARNING STYLES	EDUCE (Kelly & Tangney 2002)	✗	✓	✓	✗	Social Science	Unknown	✓	✗	✓	✗	✓	Vectors & rules	Gardner's multiple intelligence concept & types of learning goals
	Intelligent Training System on the Internet (Souto et al. 2002)	✗	✓	✗	✓	Telecommunications (Time Division Multiple Access)	Postgraduate	✓	✗	✓	✗	✓	BBN	Ross Test for Cognitive Process & the Bloom's Upper Cognitive Activities
								(related to psychological and cognitive learning styles)						

Table 3.1 New generation of Intelligent Tutoring Systems (ITSs)

	Learning Environment Characteristics													
	System	Open Learning Environment		Educational-Artificial Intelligence		Teaching Subject	Educational Level	Knowledge Represented			Reasoning / Recognising Approach		AI Approach	Underlying Theoretical Approach
		GBL environment	VLE	ITS	Intelligent Agents			Student's Cognitive Psychology Features	Student's Physical/ Biofeedback Features	Contextual Variables	Identification of Physiological Effects	Reasoning from Origin		
EMOTION	CRYSTAL ISLAND Learning Environment (Sabourin et al. 2011)	✓	✗	✓	✗	Microbiology	High school	✓	✗	✓	✗	✓	Baseline, Naive Bayes, BBNs and DBNs	Appraisal based-theory of learning emotions (Elliot and Pekrun 2007)
								(static and dynamic features related to anxiety, boredom, confusion, curiosity, excitment, frustration & concentration)						
	The Collaborative Educational Environment & JADE (Jaques et al. 2011)	✗	✓	✗	✓	JADE teached Earth time zones	HighSchool	✓	✗	✓	✗	✓	Inference rules	Belief-Desire-Intention (BDI), Ortony, Clore and Collins (OCC) model of emotion and work by Del Soldato & Du Boulay (1995)
								(related to student's goals, effort and pleasantness of events associated to satisfaction, dissapointment, joy, distress, gratitude,						
	Wayang Outpost with sensors in the classroom (Arroyo et al. 2009)	✗	✓	✗	✓	Mathematics (geometry)	High school & undergraduate	✗	✓	✓	✓	✗	Linear regression with stepwise procedure	Student self reports, Ekman theory & work by Burleson (2006)
								(related to joy, anger, surprise, fear, disgust, contempt & interest)						
	Prime Climb (Conati & Maclaren 2009)	✓	✗	✓	✗	Mathematics (Factorisation)	Undergraduate (but, it is targeted for students at Primary)	✓	✓	✓	✓	✓	Dynamic Bayesian Networks (DBNs)	Student self reports, OCC model & Big Five Personality theory
								(related to joy/distress, pride/shame & admiration/reproach)						
	LeActiveMath or LeAM (Porayska-Pomsta et al. 2008)	✗	✓	✓	✗	Mathematics (differential calculus, e.g. The chain rule)	High school & Undergraduate	✓	✗	✓	✗	✓	Machine learning techniques, decision trees & rules	Observations of actual one-to-one teaching interactions
								(performance and learning session features related to confidence, effort and interest associated with student affect)						
	Easy with Eve (Sarrafzadeh et al. 2008)	✗	✓	✓	✗	Mathematics (Part whole-addition)	Primary	✗	✓	✗	✓	✗	Artificial Neural Networks (ANNs) & Supported Vector Machine	Observations of actual one-to-one teaching interactions & Ekman & Friesen's facial coding system
								(facial gestures related to boredom, confusion, inattention & anxiety)						
	Autotutor (D'Mello et al. 2008)	✗	✓	✓	✗	Newtonian Physics, Computer Literacy & Critical Thinking	Undergraduate	✗	✓	✗	✓	✗	ANNs, Bayesian classifiers & Latent Semantics Analysis	Opinions of expert judges & Ekman & Friesen's facial coding system
								(related to boredom, engagement or flow, confusion & frustration)						
	EMASPEL (Neji & Ben Ammar 2007)	✗	✓	✗	✓	Communications Technology	Undergraduate	✗	✓	✗	✓	✗	Machine learning techniques	Ekman's (1999) basic emotions
								(facial gestures related to joy, sadness, anger, fear, disgust & surprise)						
	ERPA (Chalfoun et al. 2006)	✗	✓	✓	✗	Emotional Intelligence, Sports & other areas of general knowledge	Unknown	✓	✗	✓	✗	✓	Decision tree, ID3 algorithm & rule base	OCC model & the Eysenck Personality Questionnaire
								(attributes, personality, expectations and performance related to dissapointment, distress, joy, relief, satisfaction & fear)						

Table 3.2 Affective modelling new generation Intelligent Tutoring Systems (ITSs)

Ekman and Friesen (1978) are the most common theoretical approaches employed as a basis to interpret facial gestures.

Reasoning about students' observable behaviour from its origin, also known as Cognitive-Based Affective User Modelling (CB-AUM) (Martinho et al. 2000), employs cognitive psychological theories of emotion or affect in order to reason about the constructs involved in determining an affective state in conjunction with student characteristics, e.g. sex, and contextual features of low bandwidth such as the time elapsed in a learning session. It is noted that there are two types of affective states: moods and emotions (Ortony et al. 1990). Emotions are more dynamic in nature, i.e. change constantly and more quickly than moods, which last for a longer period of time and can also influence the experience of a specific emotional state. This approach has proven significantly effective (Jaques and Vicari 2007), but not as successful as the first approach. However, it employs feasible and low bandwidth variables. Hence, it can be employed to infer student affective state during online learning.

It is important to note that student affective states have been shown to be significantly related to contextual variables as mentioned by Arroyo et al. (2009). As in the previous approach, self-reports and expert judges are employed to record emotions over time. All this data is analysed to identify patterns and relationships to predict or classify affective states. The OCC model is the most frequently employed cognitive theory for reasoning about emotion. However, there is still debate about which should be the relevant emotions to the learning experience and therefore which emotions should be recognised. An ongoing project that it is focused on applying the CB-AUM approach is the Adaptive Learning via Intuitive/Interactive Collaborative and Emotional Systems (ALICE), which is focused on solving relevant problems of current e-learning systems through personalisation (Tu Graz et al. 2012). A sub-goal of this project is to determine the affective state of the learner in order to achieve these goals. Feidakis et al. (2011) are planning to employ self-reporting and contextual variables for deriving and testing their model.

The hybrid approach, which combines the two previous approaches, is expected to be the most effective one. However, it still has not proven very successful as can be observed in the research conducted by Conati and Maclaren (2009). The approach involves identifying the physical or physiological effects of emotion. The effects are employed in conjunction with contextual variables and factors involved in determining student affective state. The selection of these factors and the manner in which they interact to determine an emotional state is usually based on a cognitive psychological theory of emotion. As a result, it is noted that research in cognitive psychology and gaining more insight on emotion is key to apply the hybrid and the CB-AUM approaches.

3.7.2 VLEs vs. GBL environments

VLEs or GBLs have been employed by this new generation of ITSs as the context where learning happens. The emotional capacity and potential of GBL environments to communicate and affective state to the learner and to maintain it was acknowledged in Section 3.1.1. Hence, GBL environments have been employed as a source of rich interactions and information that can serve student modelling purposes. As discussed earlier, emotion is contextual and, as a result, choosing the kind of learning environment employed for teaching a determined topic or subject, its design, its instructional strategy and its characteristics are key to the kind of emotions that may arise as can be observed in the work by Conati and Maclaren (2009). Hence, in some cases, in order to identify the relevant emotions or affective states the Wizard-of-Oz technique is employed. Wizard-of-Oz is an approach employed to gain an enhanced understanding of user interaction needs and involves allowing users believe that they are interacting with an AI system when they are actually interacting with a real human. In this manner the possible answers to user interaction are unlimited (Andersson et al. 2002). However, as shown in Tables 3.1 and 3.2, ITSs have employed VLEs more frequently than GBL environments, which may be due to VLE design, which shares more similarities with actual one-to-one human tutoring. In addition, from the work by McQuiggan et al. (2008) and Arroyo and Woolf (2005), it can be observed that factors involved in determining student affective state or personal disposition are also related to the kind of subject that it is taught, e.g. mathematics, physics or social sciences. Therefore, it is important also to consider how the subject and topics that are being taught are perceived from the student viewpoint.

To combine ITSs with VLEs or GBL environments two kinds of architectures, one comprised of ITSs modules as was discussed in section 3.2 and shown in Figure 3.4, and another where the modules are implemented as agents, which comprise a multi-agent system (Wooldridge 2002) and communicate through messages to perform tasks and synchronise activities, were tested. An example of a VLE combined with an ITS and implemented with agents is EMASPEL by Neji and Ben Ammar (2007), whereas an example of a VLE combined with an ITS and implemented as a series of components or modules is LeActiveMath or LEAM by Porayska-Pomsta et al. (2008).

3.7.3 Contextual recognition variables probably related to emotion

It was observed that ITSs that attempt to be aware of students' dispositions in Table 3.1, in specific student motivation and self-efficacy, employ diverse contextual variables. In the work of MORE, discussed in section 3.3, Del Soldato (1993) focuses on creating a motivational student model, employing the contextual variables: *performance* and the *number of attempts to solve a problem* for determining little or large effort, where *performance* corresponds to the

result or outcome of the problem. For assessing whether a student's level of confidence is OK or it is low, Del Soldato (1993) utilises the *perceived level of difficulty* of a problem obtained through student self-reporting and to measure student *persistence on solving a problem* employs the following contextual variables: *number of times that the student asked for help* and *quitting without successfully solving the problem*. The latter and the *resultant performance/outcome* are also associated to student level of confidence. To know whether student independence (student achieves a successful outcome by him/herself or required help from others) is OK or low, Del Soldato (1993) uses the *number of times that the student asked for help*.

McQuiggan et al. (2008) models student *self-efficacy* in Crystal Island (see Table 3.1) employing *temporal, locational, intentional* and *physiological* interaction and contextual variables, such as *time elapsed since the question was displayed, time in the current GUI location, time on current learning goal, current cursor location, previous cursor location, number of times that the student has visited the specific location, number of problems solved, average student progression rate, average heart rate (HR) and average galvanic skin response (GSR). Temporal* variables are employed to measure student persistence. *Locational* variables are employed to know whether students are in locations where learning activities and goals are achievable. *Intentional* variables are employed to identify student (positive or negative) perceptions towards learning goals or activities and to know whether student knowledge and skills match the demands of the activity. *Physiological* variables are physiological responses acquired through biofeedback devices.

For creating a 'goals and attitudes' student model in Wayang Outpost, Arroyo and Woolf (2005) employed interaction or contextual variables. Arroyo and Woolf investigated the question of whether the student attitude related to *liking or enjoying interaction with the system*. Their results showed that this attitude is related to *the average time invested per problem*, the average of incorrect problems *where the student received help (average hints given per problem)* and *whether an improvement on student performance or progress existed* during the learning activity. In addition, to identify user attention or focus, D'Mello et al. (2008b) employs the movement of eyes, i.e. gaze. However, the mouse position or movement has also shown to be related to student concentration (Wilson 2010). In addition, there is research, especially in the area of user interaction with web content, which suggests that the mouse pointer is a suitable alternative for eye tracking for determining people's concentration or analysis behaviour (Huang et al. 2012). As a result, it is observed that all these observable variables can be related also to student emotion and may be employed to reasoning about it, since in Chapter 2, it was discussed that emotions are deeply interrelated with motivational and cognitive factors.

3.8 Summary

In this chapter OLEs and current applications corresponding to the new generation of ITSs were reviewed. From this investigation, it was observed that GBL environments and educational games have characteristics that are capable of creating and sustaining emotional communication with students. Hence, they are suitable environments and test-beds for identifying the emotions that students may experience whilst playing and learning.

ITSs are included in OLEs in order to ensure the achievement of learning goals, personalised and adaptable learning and the provision of suitable feedback. Personalised learning is the outcome of student modelling. Student modelling comprises representing knowledge for reasoning about student needs, characteristics or preferences through AI and machine learning mechanisms. Data mining techniques and PRMs are employed in order to facilitate knowledge acquisition and representation for student modelling. The KDD process comprises a series of steps such as data cleaning and data mining that can be followed in order to mine knowledge from data in repositories. ITSs initially focused on representing knowledge about students' understanding, domain knowledge and misconceptions. However, the new generation of ITSs are focused in parallel on addressing aspects of emotional intelligence that can enhance social interactions. The ultimate goal is to identify and reason about emotion. Current ITSs have employed three key approaches in order to achieve this goal: (1) identifying physical and physiological effects of emotion, (2) a cognitive-based affective user modelling (CB-AUM) approach and (3) a hybrid approach comprising the previous two. Currently, none of these approaches have proved to be reasonably effective in educational settings or in places where learning frequently takes place, e.g. online learning.

For example, the first approach has proven very successful in laboratory settings. However, a limited number of students have access to these kind of tutors and applications owing to the hardware equipment that is required as can be observed from Autotutor (D'Mello et al. 2008b) and Wayang Outpost (Arroyo et al. 2009).

The second approach, CB-AUM, employs low-bandwidth and contextual variables, which have proved highly related to emotion (Arroyo et al. 2009), and can be employed in online learning, but has not been shown to be reasonably effective (Jaques et al. 2011, Sabourin et al. 2011). In some cases, cognitive and psychological theories of emotion that have attempted to explain how emotion is produced in a context different from education have been employed (Jaques et al. 2011), such as the OCC model. However, it is not clear if these emotions are relevant to the learning-teaching experience and the theories have to be adapted to suit the demands of the educational context. In other cases, emotions and theories relevant and highly related to the learning experience have been employed, but the results expected were not obtained owing to overlooked design issues regarding the learning

environment as observed in the work by Sabourin et al. (Sabourin et al. 2011). The third approach, a hybrid approach combining the previous two approaches, is expected to be the most successful (Sarrafzadeh et al. 2008). However, presently results have not shown significant promise (Conati and Maclaren 2009). This hybrid approach has inherited the strengths but also the weakness of the other two approaches.

These problems signal an opportunity to conduct research within affective student modelling through the creation of an emotional student model that can be employed in online learning and reason about student observable behaviour with low bandwidth and feasible variables. The Control-value theory of achievement emotions by Pekrun et al. (2007), reviewed in Chapter 2, has not been employed previously for creating a computational and emotional student model. It shows that this goal is feasible and worthy of investigation, since GUIs that do not respond intelligently or suitably and address emotion are considered untrustworthy, inept and may limit student performance. The intelligent behaviour of ITSs relies on the student model, which leaves them capable of understanding and reasoning about student behaviour, characteristics, needs or preferences. As a result, to create intelligent adaptable learning environments, it is necessary to provide them with capabilities to reason about emotion. The following chapter discusses our computational model of student emotions employing Control-value theory for reasoning about student emotions in an online GBL environment.

Chapter 4: A Computational Model of Student Emotions

This chapter discusses a proposed computational model of student emotion for reasoning about student *achievement emotions* whilst interacting in a GBL environment. The Control-value theory (Pekrun et al. 2007) forms the basis for deriving this model. This model employs a dynamic sequence of BBNs for representation and contextual variables (e.g. mouse focus and requests for help), answers to questions in-game dialogues and physiological variables, i.e. Galvanic Skin Reponse (GSR) signals, for recognition.

The main hypothesis of this dissertation is that: an emotional student model based on Control-value theory (Pekrun et al. 2007), using only contextual variables, will reason about relevant emotions in an online GBL environment with the required accuracy. A subsidiary hypothesis of this dissertation is that: the accuracy of such an emotional student model can be enhanced by adding physiological variables. These hypotheses lead to the following questions:

1. Which are the factors or variables to be considered when reasoning about student emotion in an online GBL environment?

2. Is there mutual causation between the antecedents and effects of *achievement emotions* over time?

3. What is the best approach for the derivation of the dynamic sequence of BBNs? And how can the contribution of the factors/variables be evaluated in order to determine if they are relevant to the classification?

4. Does incorporation of physiological variables, i.e. Galvanic Skin Response (GSR), increase the accuracy of classification within the model?

5. Does a qualitative and quantitative evaluation of the GBL environment, in which our emotional student model resides, help to achieve an enhanced understanding of student experience of *achievement emotions* and the factors involved in inferring them?

4.1 Approach to reasoning about emotion

A key problem in creating an emotional student model is determining the relevant emotions. Our emotional student model focuses on modelling the emotions relevant to student learning in online educational gaming - *achievement emotions* (Pekrun et al. 2007). We chose to pursue an online solution since, as discussed in Chapter 3, a larger student population can be accessed and students feel more comfortable, and are frequently accustomed to, receiving education online. Hence, this investigation focuses principally on using contextual and low-bandwidth variables to classify student emotions. The main approach that we propose to follow to attain a representation using only contextual variables and Control-value theory, is 'reasoning about emotion from its origin', also known as Cognitive-Based Affective User Modelling (CB-AUM). When physiological variables are incorporated, the approach employed is a Hybrid of CB-AUM and identifying the effects of emotion. We propose to create the emotional student model using a dynamic sequence of BBNs since, as noted in Chapter 3, this approach selects BBNs from a library according to the state of the System and BBNs include similarities with DBNs, which have proven more successful than other AI techniques when reasoning about emotion using contextual variables (Sabourin et al. 2011).

Also in Chapter 3 it was observed that to achieve a suitable representation using Bayesian networks, the steps involved are: (1) the selection of relevant factors or input recognition variables, (2) determining the network structure using the relevant variables and (3) the parameter learning - defining the strength of the associations between random variables. The defined model should then be evaluated in its performance for reasoning about student emotion, i.e. propagating the evidence using a learning algorithm such as max-propagation. The method employed for evaluating the model depends on the availability of data. This section explains in detail the process followed to achieve the complete representation of our emotional student model and the assessment of its performance whilst recognising student achievement emotions.

4.2 Recognition of emotions

In order to identify at a high-level the variables involved in determining the experience of *achievement emotions*, first, we focus on examining the Achievement Emotions Questionnaire (AEQ) (Pekrun et al. 2005). The AEQ comprises affective, cognitive, motivational and physiological question items, i.e. factors, which are assessed *before*, *during* or *after* a specific activity, i.e. during a class or lecture, carrying out independent study or taking a test. In the AEQ, students select from a likert-scale (from 1 to 5) the degree to which they agree or disagree with a statement, where 1 corresponds to 'Strongly disagree' and 5 corresponds to 'Strongly agree'.

In our case, we decided to focus on the factors employed to assess student emotions whilst carrying out independent study, since it is expected that students interact with the GBL environment for this purpose. Also, we decided to focus on cognitive and motivational factors for online educational gaming. In the case of on-site educational gaming, we also incorporate physiological variables. It is important to highlight that when considering both contextual and physiological variables, the approach moves from CB-AUM to a hybrid approach. We decided not to consider affective items, since after analysing the AEQ, it was observed that these correspond to student self-reporting of the emotional state that students may be experiencing. For example, to identify whether the student feels anger before conducting independent studying/learning, the affective item employed is: 'I get angry when I have to study' (Pekrun et al. 2005, p. 20) and for determining if the student feels anxiety, the affective item employed is: 'When I look at the books I still have to read, I get anxious' (Pekrun et al. 2005, p. 21).

We analysed the AEQ items in Pekrun et al. (2005), in order to identify the key motivational and cognitive factors relating to control or value appraisals for determining *achievement emotions* in the time frames *before*, *during* and *after* a *learning-related* activity, such as conducting independent study. We created a summary of these factors, shown in Figure 4.1. The manner in which these factors were inferred from the AEQ items is exemplified in Table 4.1. It can be observed that the factors employed to determine the specific emotion change depends on the time frames – *before*, *during* or *after* - and the object, e.g. outcome or activity.

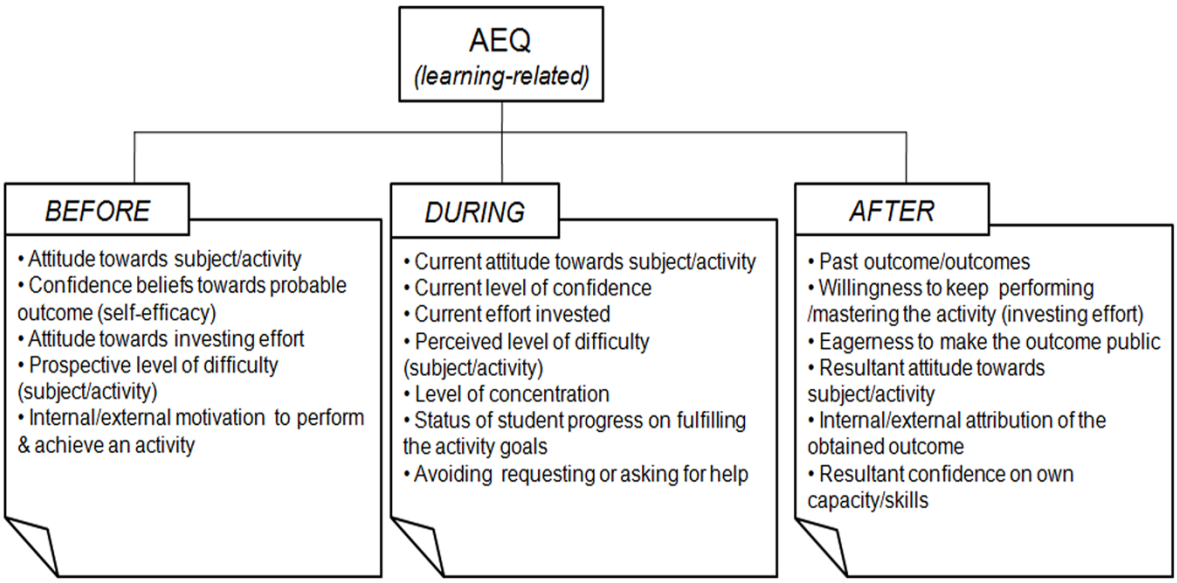


Figure 4.1 Summary of factors on the AEQ for assessing *achievement emotions*

Item as stated in the AEQ	*Factor(s) inferred*	*Time frame*			*Type of factor*	
		Before	*During*	*After*	*Cognitive*	*Motivational*
'I feel confident that I will be able to master the material' (Pekrun et al. 2005, p. 18)	Confidence towards possible outcome (self-efficacy)	✓			✓	
'Because I get so upset over the amount of material, I do not want to begin to study'	Attitude towards investing effort	✓				✓
'I worry whether I'm able to cope with all my work' (Pekrun et al. 2005, p. 21)	Current level of confidence and perceived level of difficulty		✓		✓	
'I study more than is required because I enjoy it so much' (Pekrun et al. 2005, p. 17)	Willingness to keep performing/mastering the activity (investing effort)		✓			✓
'I think I can be proud of my accomplishments at studying' (Pekrun et al. 2005, p. 19)	Past outcome			✓	✓	
'Because I have had so much trouble with the course material, I avoid discussing it' (Pekrun et al. 2005, p. 22)	Internal/external attribution of the attributed outcome and resultant confidence on own capacity/skills			✓		✓

Table 4.1 Procedure used to identify key factors for determining *achievement emotions*

4.2.1 Variables to model prospective-outcome emotions

From the analysis conducted to identify the factors employed in the AEQ and summarised in Table 4.1, it was concluded that the variables employed for identifying the student achievement emotions, i.e. prospective outcome emotions, in the time frame before, are student attitudes and beliefs related to the future outcome of the learning activity. An attitude is a learned affective or evaluative pre-disposition to a object, the residue of past experiences, which is employed by a person to respond consistently or act accordingly, e.g. positively (for) or negatively (against) (Ajzen 2005). It is frequently associated with a choice between alternatives. Also in comparison to personality traits, attitudes are more dynamic, i.e they can change over time. A belief is information about an object (Ajzen 2005), e.g. attributes. Variables that we selected to be used to identify *prospective outcome emotions* are as follows:

- *Attitude towards the subject or the activity*: the student's positive or negative pre-disposition towards learning a subject or performing a learning activity. It cannot be observed directly. Hence, it is inferred from consistent behaviour (Ajzen 2005).
- *Confidence beliefs towards the probable outcome:* These beliefs correspond to the term *self-efficacy* introduced by Bandura (1995), which is related to a person's beliefs in his/her own capabilities that influence their response to, and perceived control over, a specific activity or situation.
- *Attitude towards investing effort:* corresponds to the student pre-disposition, for or against, devoting effort to the learning activity, which is related to student's motivation (Keller 1983).

- *Prospective level of difficulty*: relates to student self-efficacy and corresponds to perceived control over the learning activity, subject or situation.
- *Internal/external motivation to perform or achieve an activity:* relates to the source of student intention to learn a subject or perform a learning activity. If the person's intention originated from the person by himself/herself (internal) or it was imposed, since it is desired by someone else (external). According to Ajzen (2005), an intention or behavioural intention is a particular category of beliefs in which a behaviour (information) is known about a person (object). Intention is related to a person's attitudes and its strength can be measured as the likelihood of a person to act or behave in a specific manner.
- Past outcome (pre-test mark): the latter outcome that the student just achieved may assist him/her to reflect upon his/her own capabilities.

The student *gender* is also included as another variable capable of influencing specific achievement emotions in a particular group of the population. In order to enquire about student beliefs as part of the interaction offered by the GBL environments, cut-scenes, comprising game dialogues, can be employed. As stated in Chapter 3, cut-scenes commonly provide the introduction and conclusion of a game or present narrative (Collins 2008). The latter can reward the player for his/her accomplishments. The student has to be asked to self-report his/her emotional state at the end of the cut-scene. The objective is to discover the student *achievement emotion* related to the state of these variables in order to derive, train and evaluate a classifier created using Artificial Intelligence (AI) techniques. As discussed in Chapter 3, this method of questioning and encouraging self-reflection has been successfully employed by work attempting to identify the student motivational state (De Vicente 2003, Del Soldato 1993), self-efficacy (McQuiggan et al. 2008) and emotion (Conati and Maclaren 2009, Sabourin et al. 2011)

4.2.2 Variables to model activity emotions

Regarding the time frame during, the variables employed by Pekrun et al. (2005) are related to the learning activity that the student is performing and the progress that he/she is making. However, in this case, it is not feasible to ask students to frequently self-report the state of all these variables in conjunction with their emotional state while focusing on solving the game challenge(s), since it is probable that this procedure will be time consuming and lead to significant deviation of student attention from the learning goal. Hence, we decided to employ interaction or context variables, which have been shown in the work reviewed in Chapter 3 to be significant regressors of the key factors employed by Pekrun et al. (2005), i.e. it is assumed that the effect of these key factors is represented by interaction or context variables, which in turn are presumed to be related to student behaviour patterns, i.e. student actions, associated with student achievement emotion. The observable variables selected come from

the related work in Chapter 3 and have been used to identify student motivation (Del Soldato 1993), self-efficacy (McQuiggan et al. 2008), goals and attitudes (Arroyo and Woolf 2005) and concentration (Wilson 2010). We propose that these observable behaviour variables can assess student control and value with the required accuracy. Using as a reference the categorisation of observable variables by McQuiggan et al.(2008), also discussed in Chapter 3, the selected temporal, intentional and locational interaction/contextual variables employed to assess student achievement emotions while interacting with an online GBL environment are shown in Table 4.2. The locational variables are specifically employed in this dissertation to evaluate student concentration or attention.

Type of variable	*Variable*	*Description*	*Associated factor(s) of control or value*				
			Effort	*Confidence*	*Perceived level of difficulty*	*Attitude towards the activity*	*Concentration*
Temporal	Interval of interaction	The total time that the student has interacted, since the game challenge is started	✓				
	Time to achieve the learning goal(s)	The time that the student invested for achieving the learning goal the first time	✓	✓	✓		
Intentional	Outcome	The result that it most likely to be achieved and directly associated to student performance		✓	✓	✓	
	Times asked help	The number of times that the student asked for help	✓	✓			
	Attempts alone	The number of attempts by the student to solve the challenge alone (without help)	✓	✓			
	Estimated independence	Results from the difference between the number of attempts alone and the number of times that the student asked for help	✓	✓			
	Overall attempts	The total number of student attempts with and without help	✓	✓		✓	
	Average quality of tutoring feedback	The average value calculated from the student qualitative evaluation related to how useful, he/she finds the help or instruction provided	✓		✓	✓	
	Type of outcome or end condition	Indicates whether the student obtained a successful outcome, committed a misconception or quit the game challenge		✓	✓	✓	
Locational	Focus coarse value	The average value of the mouse position on the screen associated to student attention	✓			✓	✓

Table 4.2 Summary of interaction variables selected for modeling activity emotions

In the case where the student is interacting with an on-site GBL environment, physiological variables can be also employed in order to enhance the classification accuracy of the model, since as discussed in Chapter 2, physiological changes are associated with a person's experience of emotion, so we decided to also conduct a test of the GBL environment on-site and acquire student GSR signal. Pekrun et al. (2005) include items to enquire about the participant heart rate (HR). However, in Chapter 2 we discussed that GSR is more sensitive to,

and is significantly associated with, changes in student emotion. Therefore, in order to acquire student GSR signal in real-time, we suggest creating a biofeedback device (see Chapter 2) using the homemade GSR signal reader built by Gasperi (2010) using the LEGO NXT brick. However, we propose to modify the software to connect via Bluetooth and acquire real-time data, i.e. the raw values of students GSR signal.

4.2.3 Variables to model retrospective outcome emotions

According to Control-value theory, at the time frame *after*, i.e. end of the learning activity, the student focuses on the outcome achieved and presented. Hence, his/her emotion is the product of reflection on past performance and progress. As a result, the latest state of the interaction variables defined in Table 4.2 can be employed, such as the latest outcome showed to the student, the latest state of the estimated independence (attribution of the result) and the type of outcome or end condition (willingness to keep interacting). However, in addition to these variables the intentional variable: *publishing outcome* should be considered, which is related to student enthusiasm for publicising the outcome achieved.

4.3 Representation of emotions

In Chapter 2, it was noted that Pekrun et al. (2007) discuss that there is mutual causation between antecedents and effects of *achievement emotions* over time. Their consideration of the dimension *time frame* and *object* for defining specific types of achievement emotions suggest that the observable/interaction variables evolve over time and that there are temporal relationships that need to be appropriately represented. Also in Chapter 3 it was established that emotional student models that have modelled temporal relationships of emotions over time more appropriately and accurately have employed this approach (Conati and Maclaren 2009, Sabourin et al. 2011). It is also important to underline that, according to the theory by Pekrun et al. (2007) reviewed in Chapter 2, *control* and *value* are categorical or nominal variables with two or more categories, where *control* can be classified as ‘high’, ‘medium’, ‘low’, ‘irrelevant’, ‘self’ or ‘other’ and *value* can be classified as ‘positive’, ‘none’ or ‘negative’. Bayesian networks can be comprised of categorical nodes. Hence, we decided to deploy a dynamic sequence of BBNs for representing student emotion.

As discussed in Chapter 3, there are two steps involved in acquiring a Bayesian model from data: (1) obtaining the network structure and (2) determining the parameters, i.e. finding relationships between random variables and the strength of these associations. The methodology that we intend to employ in this dissertation to achieve these two objectives is shown in Figure 4.2.

Figure 4.2 Methodology used to define a dynamic sequence of BBNs

First, we employ Probabilistic Relational Models (PRMs) to facilitate the understanding of the involved domain in order to define a preliminary generic emotional model and identify the random variables that may be included in our emotional student model. Then with the data acquired from student interaction with the GBL environment, we analyse the Pearson correlations, i.e. potential associations and their strength, and then we employ Binary or Multinomial Logistic Regression (MLR) to determine which of these variables may be the best predictors for control and value. The main objective of obtaining the Pearson correlations and applying MLR is to achieve more information of the affective domain. Finally to complete the definition of Bayesian networks, the Necessary Path Condition algorithm is applied and uncertain relations are solved using the information acquired from the domain using Pearson correlations and Binary/MLR. The Expectation Maximisation (EM) is employed to learn the network parameters. This section provides more details about the rationale behind this approach and how each of the steps involved was applied.

4.3.1 Defining a generic model of student emotion

As established in Chapter 3, to facilitate the derivation of Bayesian models, i.e. reducing the complexity and time invested acquiring domain knowledge, Probabilistic Relational Models (PRMs) are employed (Sucar and Noguez 2008). PRMs also ease the effort of extending the model to other domains, i.e. derivation of generic models. As a result, in order to achieve these advantages, a PRMs approach is taken to derive three BBNs, which include historic nodes and relations of the type $\Pr(X_t \mid X_{t-1})$, one per each type of *achievement emotion*: *prospective outcome*, *activity* and *retrospective outcome* emotions. A generic PRM of Control-value theory for reasoning about student *achievement emotions* at a high level is shown in Figure 4.3. The *Student* class comprises features such as *student gender* or *group* to which the student belongs. The student *exploration/interaction* class represents a game dialogue or a game challenge, which has attributes, i.e. answers to questions in game dialogues or features related to student observable behaviour, for reasoning about student *achievement emotions*. *Time frame* indicates a temporal slice, which may correspond to the period of time *before (Figure 4.4)*, *during (Figure 4.5)* or *after (Figure 4.6)* an activity, i.e. game challenge.

The time frame *before* involves a cut-scene or game-dialogue previous to each game challenge, the time frame *during* involves the student playing and learning with the game challenge and the time frame *after* corresponds to the notification of the achieved game/learning outcome. The *Focus of emotion class* corresponds to the key feature available to assess student cognition, which can be the future or past outcomes of the game challenge or the ongoing student progress. The generic PRMs for each specific time frame (Figures 4.4 to 4.6) are derived from the high-level and generic Control-value PRM in Figure 4.3.

Figure 4.4 shows the preliminary skeleton of the generic PRM designed for reasoning about *prospective outcome emotions* in cut-scenes comprised of game dialogues that in turn include motivational and cognitive question items adapted to the storytelling and focused on gathering student data through their direct answers. At this point, the dialogue must focus student attention on the future outcome of the activity to be performed and encourages him/her to reflect on his/her own attitudes, beliefs, skills and capabilities in order to respond to questions from a personal viewpoint. If the student assumes the role of the player character as a reference for answering, this can add uncertainty to the emotional student model. In this case, the random variables (answers) may or may not be specifically related to *control* and *value* (parent nodes). The relationships between these nodes and their likely children nodes (random variables) are indicated with dotted lines in Figure 4.4. Also, it is not certain whether child nodes are related to each other and these relationships are signalled with dotted lines. *Control* and *value* nodes at time *t-1* come from the *retrospective outcome* emotions PRM, while *control* and *value* at time *t+1* correspond to the *activity* emotions PRM.

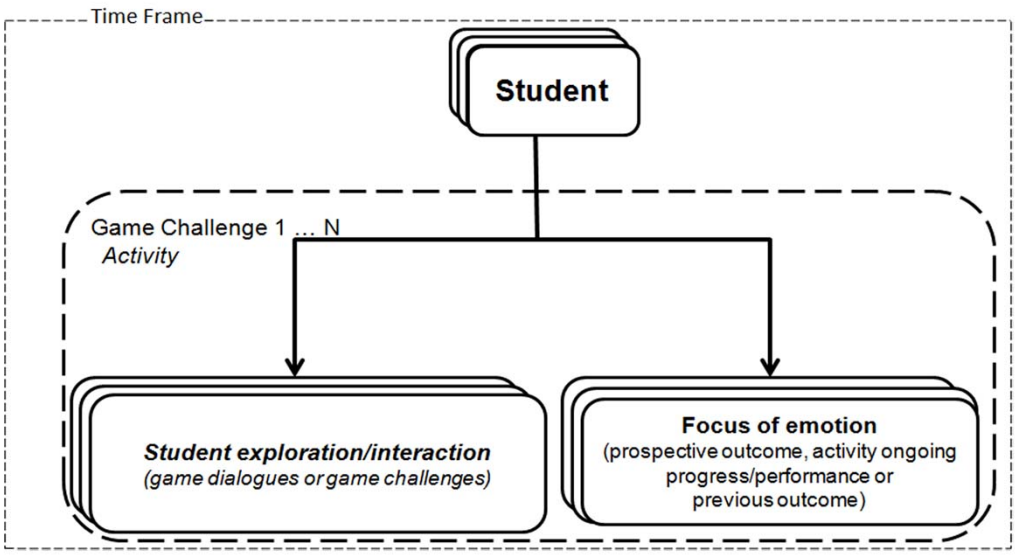


Figure 4.3 Generic Control-value PRM

Figure 4.5 shows the preliminary skeleton of the generic PRM corresponding to *activity emotions*. At this stage the student focuses on the progress being completed in the per-

formed game challenge. The available random variables for reasoning about student *achievement emotions* relate to student interaction with, and exploration of, the online GBL environment, e.g. *intentional*, *temporal* and *locational* observable variables.

In cases where the student is interacting with an on-site GBL environment, physiological random variables will also be available. However, at this point it is not clear which variables are related to *control* and *value* or if a relationship exists at all, so these relationships are indicated with dotted lines in Figure 4.5 as the relationships between observable random variables. *Value* at time t-1 and *control* at time t-1 may come from the prospective outcome, activity and retrospective outcome emotions BBNs. Figure 4.6 shows the preliminary skeleton of the generic PRM corresponding to student *retrospective outcome emotions*. Similar to the generic PRM corresponding to student *activity emotions*, intentional, temporal, locational and possibly physiological random variables are employed for reasoning about student achievement emotions. The availability of random variables depends again on the kind of GBL environment in which the student is interacting. However, special attention is given to the latest state of these variables and past outcomes, but also in this case, it is not assured that the potential random variables are related to *control* and *value*, or whether these variables are related to each other (associations indicated with doted lines). In this case *control* and *value* at t-1 come from the generic *activity emotions* PRM, while *control* and *value* at t+1 correspond to the generic *prospective outcome emotions* PRM.

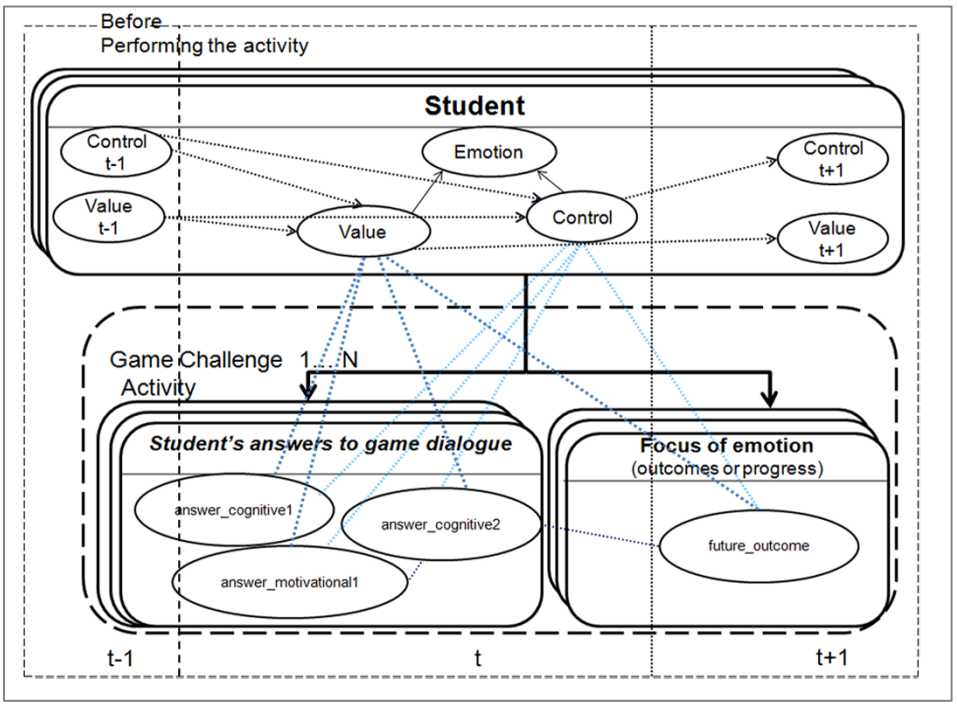


Figure 4.4 *Prospective outcome emotions* generic PRM

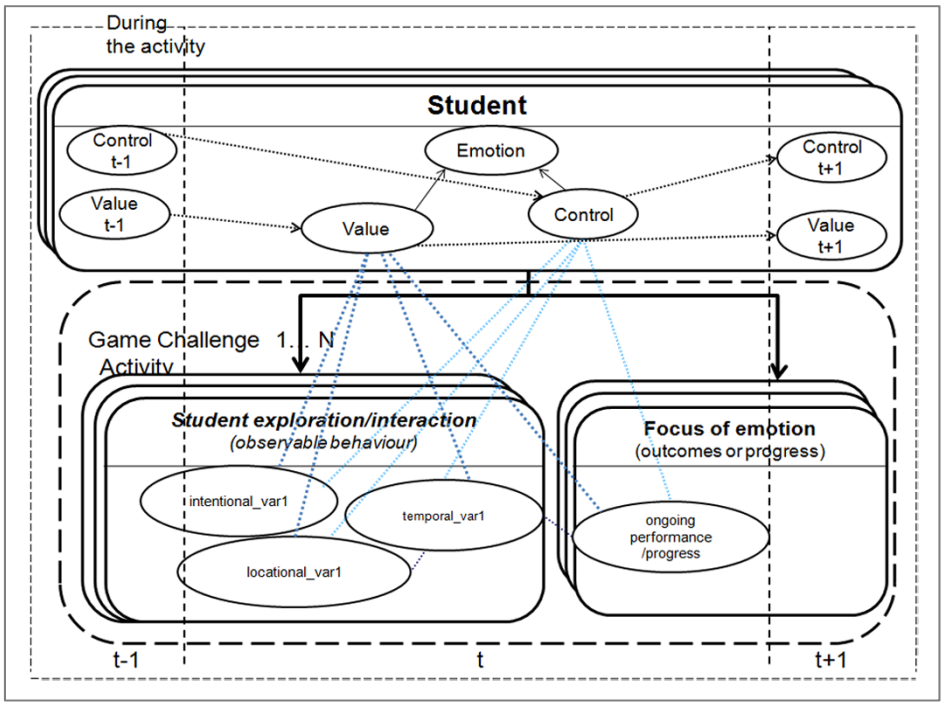


Figure 4.5 *Activity emotions* generic PRM

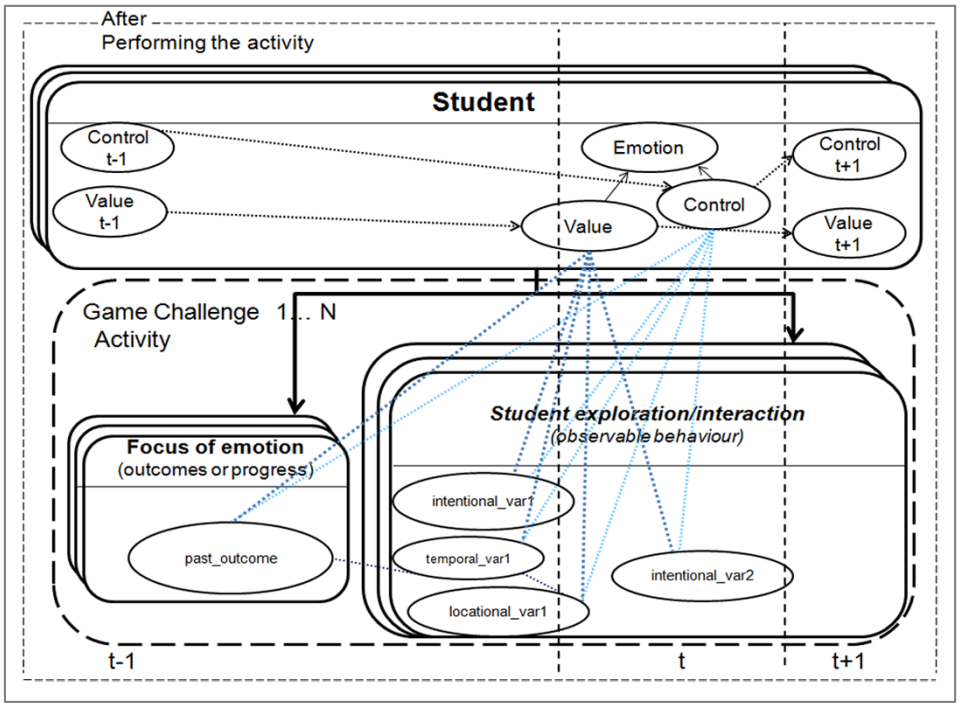


Figure 4.6 *Retrospective outcome emotions* generic PRM

4.3.2 Instantiating and completing the dynamic sequence of BBNs

We propose that the derived generic emotional model will be employed as a reference for the kind of variables that are available in the specific GBL environment context and that may be included. This model has to be instantiated according to this GBL environment. We then face the problem associated with identifying whether the potential random variables are definitely

related to *control* and *value* or whether they are interrelated, i.e. they define the acyclic structure of the networks. To determine these associations with certainty, the student interacts with the specific GBL environment and data corresponding to the potential random variables is recorded. As stated in Chapter 3, data mining techniques can be employed to reduce the uncertainty of the data, unveil data patterns and discretise random variables.

Subsequently, we intend to obtain Pearson correlations to identify associations and their strengths. Arroyo and Woolf (2005) employed them as references to create their DBN structure as described in Chapter 3. Later, we plan to apply Binary or MLR, depending on the number of categories that comprise the dependent variable, *control* or *value*, in order to identify from all the identified associations which variables are the most meaningful. The rationale behind this is that Bayesian models are a kind of logistic regression, as mentioned in Chapter 3. We employ all the available data to apply the Necessary-Path Condition algorithm, since our data is scarce, to obtain the networks structure. However, to solve the uncertain associations, we employ all the information obtained through Pearson correlations and the results of applying Binary/MLR. Once the structure is completed, we will proceed to apply the EM learning algorithm to learn the Conditional Probability Tables (CPTs).

4.3.3 Reasoning about emotion: Measuring performance

To reason about emotion this model propagates evidence using the max-propagation algorithm discussed in Chapter 3. However, to determine how the model performs over fresh data, we will employ stratified 10-fold cross-validation, which has shown to be quite reliable and is highly adopted by related research discussed in Chapter 3. With this method, 90% of the data will be retained for training and 10% of the data will be employed for testing. This procedure is executed 10 times and each time the model is evaluated using data that has not been used to create the networks structure nor for defining the probabilities in the CPTs. Consequently, we will obtain one BBN per training data set. However, the BBN that we propose to report is the BBN created using all the available data and we will obtain an estimate of the performance of this network over fresh data.

4.4 Experience of achievement emotions in GBL environments

We discussed in Chapter 3 that the experience of users in emotional computer applications are appraised with quantitative-scientific and open-ended interpretation techniques (qualitative approach). In this dissertation, to obtain an enhanced understanding of student viewpoints or experiences related to student achievement emotions and to assess the performance of the GBL environments, these two approaches are also employed. We believe that these methods can signal other observable variables or student features that can be employed in future work to enhance the performance of our proposed emotional student model.

The quantitative method employed focuses on investigating whether the GBL environment aided the student in achieving enhanced knowledge and understanding. The qualitative approach allows insight into specific elements of the emotional and educational gaming experience, including which elements influenced whether the student did or did not achieve the learning goals. Also, these elements may contribute to making the experience enjoyable or not from the student perspective. In this case we assess the appropriateness of fantasy, challenge and ease of interaction offered by the GBL environment. As mentioned in Chapter 3, these elements are signalled by Malone (1981, 1984) for designing enjoyable GBL environments. Also, student comments and suggestions will be encouraged as a resource for open-ended interpretation of student learning and gaming experience. In this manner, in addition to obtaining a global student perspective, we obtain particular viewpoints of the quality of the interaction offered by the GBL environment.

4.5 Summary

This chapter describes our approach to create a model of student emotion based on Control-value theory (Pekrun et al.). This emotional student model is mainly designed to be employed during online educational gaming using questions to answers in game dialogues and *intentional, temporal and locational* observable behaviour variables for reasoning about emotion. However, our proposed emotional student model can also be employed in on-site GBL environments, where *physiological* variables – specifically GSR signal - can be employed and may enhance emotion classification accuracy, i.e. recognition. The procedure employed for selecting and identifying the features or random variables from Control-value theory is explained. Bayesian models, e.g. a dynamic sequence of BBNs, are employed for representation of student emotions. The creation of these networks requires two stages to be defined: (1) learning the network structure and (2) learning the parameters. Here, we explore whether we can employ Necessary-Path Condition in combination with Person correlations and Binary/Multinomial Logistic Regression (MLR) to select relevant variables and solve uncertain relations in order to define the network structure.

PRMs facilitate the derivation of Bayesian models by providing an intuitive representation of the knowledge domain deployed in our emotional model. Three generic PRMs were defined, each corresponding to a type of *achievement emotion*: *prospective outcome, activity and retrospective outcome emotions*. EM is an algorithm that will be employed here to define network parameters: conditional probabilities. For reasoning or recognising emotion, i.e. propagate evidence, we will employ the max-propagation algorithm. For determining how accurate the model is in its reasoning/recognition over fresh data cross-validation is employed. In the following chapter, we discuss the design of PlayPhysics, an emotional game-based learning (GBL) for teaching physics, in which our computational model will be incorporated.

Chapter 5: PlayPhysics Design

This chapter discusses the design of PlayPhysics, an emotional game-based learning (GBL) environment for teaching physics. PlayPhysics will incorporate our emotional student model and enable testing of the two hypotheses discussed in chapter 4. This chapter discusses the requirements analysis and PlayPhysics design. PlayPhysics covers physics topics that are generally considered the most challenging in an introductory physics course at undergraduate level. PlayPhysics is implemented with the Olympia architecture and software platform, a generic platform that combines the functionality of Intelligent Tutoring Systems (ITSs) and Game-based learning (GBL) environments. The means by which our emotional student model is instantiated within PlayPhysics is also discussed here. PlayPhysics also includes the learning companion, M8-robot, which provides directional instruction, mirrors self-reported student emotion and in appropriate cases encourages and praises the student. In addition, PlayPhysics integrates capabilities for encouraging students to self-report their emotional state and also utilise when appropriate, student physiological or biofeedback data, e.g. Galvanic Skin Response (GSR).

5.1 User requirements analysis

In order to effectively design GBL environments a key step is to conduct a pre-analysis phase as suggested by Akilli and Cagiltay (2006) in their FIDGE (Fuzzified Instructional Design Development of Game-like Environments) model. In this phase, a prospective target population is selected for the chosen subject to be taught. Then an investigation must be conducted in order to determine if the population and the subject are suited to teaching with this method. Once the population and the subject are selected, the next step is to set the learning goals that need to be achieved. An important step in this phase is to consider student preferences and needs for learning and lecturer instructional strategies, suggestions and comments, which are captured through interviews or surveys. Additionally, different game genres are examined and several development tools reviewed in order to choose the most suited. The selected group is mainly comprised of students at undergraduate level pursuing an Engineering or Science degree. Also, students undertaking their last year of high school and wishing to pursue an Engineering or Science degree may be targeted, since it was decided to focus on teaching introductory level physics. Students were chosen in this

age range, since the majority are already 18 years old or over and are considered adults. As a result, the ethical requirements are not as rigorous as when working with a sensitive population, e.g. children or older people. The subject of physics was chosen for four key reasons:

1. Physics is considered a challenging and complex subject (Er and Dag 2009).
2. There is a recognised challenge in maintaining student engagement while teaching physics, since its underlying structure can be difficult to grasp (Muñoz et al. 2009a).
3. Education requires an improvement, especially in the UK, and experts suggests that such an enhancement may be achieved by having an effective foundation in Science, Technology, Engineering and Mathematics (STEM) (Khan 2011).
4. Physics is a subject where students can learn by doing, i.e. problem-based learning, since Physicists interrogate and understand nature through experimentation.

In order to gain awareness of student preferences, characteristics and needs whilst learning introductory physics and also to consider lecturer comments and suggestions, two online surveys, i.e. one for lecturers and one for students, were conducted between September and October 2009. Lecturers and students from two high level institutions, namely Trinity College Dublin (TCD) and Tecnológico de Monterrey Mexico City (ITESM-CCM) participated in this survey. In total, four lecturers and fifty-three students participated in these surveys. From Trinity College Dublin, one lecturer and eighteen students participated, while from ITESM-CCM, three lecturers and thirty-five students participated. Two versions of each online survey questionnaire were created: Spanish and English. Appendix A shows the English and Spanish versions of these questionnaires. The data gathered from both surveys was analysed with Statistical Package for the Social Sciences (SPSS) version 17 (IBM 2012).

5.1.1 Student user requirements questionnaire results

The participants at ITESM-CCM were composed of 17 males and 18 females in an age range from 18 to 23 all of whom are pursuing an Engineering related degree, such as Mechanical and Electronic Engineering or Biotechnology. The participants at Trinity College Dublin (TCD) were composed of 11 males and 7 females in the same age range as students at ITESM-CCM, but all the TCD students are pursuing a Science related degree, such as Physics of Motion, Physics and Chemistry of Advanced Materials or Natural Sciences. When students in both institutions were asked about their preferences for learning from auditory, non-verbal, tactile or kinaesthetic and verbal stimuli, (Figures 5.1 and 5.2), results showed that students have a preference for learning by doing (tactile or kinaesthetic), since between 55% and 56% of the students reported feeling very comfortable with this strategy, which is compatible with the skills for studying engineering or a degree in Science (see Figures 5.1

and 5.2). However, between 43% and 44% reported feeling uncomfortable with auditory stimulus for learning. In the results for verbal stimulus and non-verbal stimulus (see Appendix A). The number of students reporting feeling comfortable and fairly comfortable with these stimuli was respectively between 20% and 30% in each category.

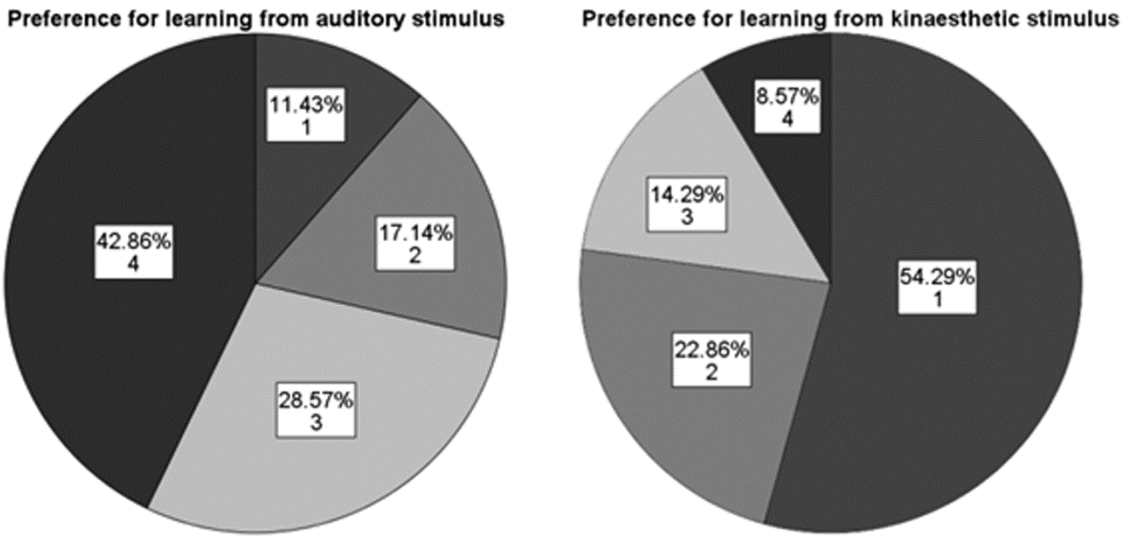


Figure 5.1 ITESM-CCM student preferences for learning from stimuli

From the answers to question 6 of the student questionnaire (see Appendix A), the respondents' Myers-Briggs personality types were inferred. Personality type informs preferred student play style according to the DGD1 model, discussed in Chapter 3. Relationships between student personality types and play styles of the DGD1 model are show in Tables 5.1 and 5.2. As shown in table 5.1 it can be observed that eleven students from ITESM-CCM, equivalent to 31.43% of the total participants, are categorised as Extraversion, Sensing, Thinking and Judging (ESTJ), whereas 9 students (25.71% of the participants) are categorised as Extraversion, Sensing, Thinking and Perceiving (ESTP). These two types are present in approximately 60% of the entire population of students. Both types direct their lives using the external world as their main reference (Bayne 1997). They are social, conversational and confident. In addition both personality types have a secondary mode or preference for dealing with things internally using logical and rational thinking. However, ESTJs are critical and firm with people – they are frank but sometimes lack tact. They are also organised and like to accomplish closure of activities and projects. ESTPs are straightforward risk takers who take action and give less importance to introspection. They are practical, observant and spontaneous and like truly thrilling experiences.

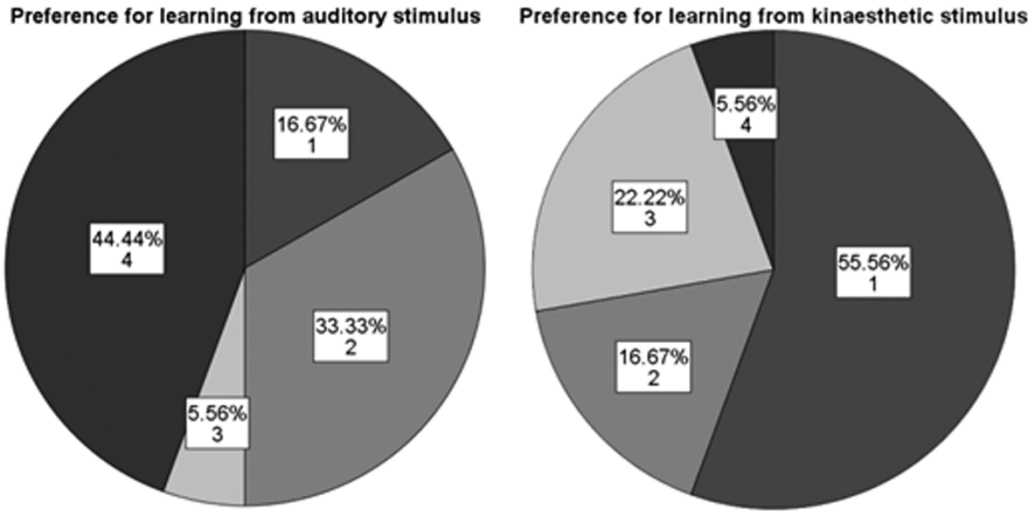


Figure 5.2 TCD student preferences for learning from stimuli

			Myers-Briggs personality type											
			ESFP	*ESFJ*	*ESTP*	*ESTJ*	*INFP*	*INTP*	*INTJ*	*ISFP*	*ISFJ*	*ISTP*	*ISTJ*	*Total*
DGD1 play style	*Conqueror*		0	0	0	11	0	0	1	0	0	0	3	15
		% of Total	.0	.0	.0	31.4	.0	.0	2.9	.0	.0	.0	8.6	42.9
	Participant		0	2	0	0	0	0	0	0	1	0	0	3
		% of Total	.0	5.7	.0	.0	.0	.0	.0	.0	2.9	.0	.0	8.6
	Manager		0	0	9	0	0	1	0	0	0	2	0	12
		% of Total	.0%	.0	25.7	.0	.0	2.9	.0	.0	.0	5.7	.0	34.3
	Wanderer		2	0	0	0	2	0	0	1	0	0	0	5
		% of Total	5.7	.0	.0	.0	5.7	.0	.0	2.9	.0	.0	.0	14.3
Total		*Num. Students*	2	2	9	11	2	1	1	1	1	2	3	35
		% of Total	5.7	5.7	25.7	31.4	5.7	2.9	2.9	2.9	2.9	5.7	8.6	100.0

Table 5.1 ITESM-CCM student DGD1 play styles for Myers-Briggs personality types

As shown in table 5.2, 22% of students at Trinity College Dublin correspond to an ESTJ type, which was also identified as a frequent personality type for students at ITESM-CCM. However, 22.22% of students have an Extraversion, Sensing, Feeling and Judging (ESFJ) type and 16.7% of students have an Introverted, Sensing, Thinking, Judging (ISTJ) type. These personality types account for approximately 61% of the participants. ESFJs, like ESTJs, are also focused on the external world, but their interaction is driven by their feelings. They are interested in people and life. As a result, they use their senses to collect comprehensive information about people. This skill means they are particularly effective at understanding and evaluating people. From their perspective, security, stability, tradition and ap-

proval from others are extremely important. As a result, they measure their values and moral norms against society. They may be insecure, over sensitive, extremely controlling and focus on pleasing others, in addition to being cooperative, practical, thorough, consistent, organised, tactful, helpful and enthusiastic.

			Myers-Briggs personality type								
			ENFJ	*ENTP*	*ENTJ*	*ESFJ*	*ESTJ*	*INFJ*	*INTJ*	*ISTJ*	*Total*
DGD1play style	*Conqueror*		0	0	2	0	4	0	2	3	11
		% of Total	.0	.0	11.1	.0	22.2	.0	11.1	16.7	61.1
	Participant		1	0	0	4	0	1	0	0	6
		% of Total	5.6	.0	.0	22.2	.0	5.6	.0	.0	33.3
	Manager		0	1	0	0	0	0	0	0	1
		% of Total	.0	5.6	.0	.0	.0	.0	.0	.0	5.6
Total		*Num. Students*	1	1	2	4	4	1	2	3	18
		% of Total	5.6	5.6	11.1	22.2	22.2	5.6	11.1%	16.7%	100.0

Table 5.2 TCD student DGD1 model play styles for Myers-Briggs personality types

ISTJs focus mainly on their inner world and have a strong and internal sense of duty that drives them to be organised and methodical and encourages them to achieve tasks and goals. They also deal with the external world rationally and logically. Security, peace, loyalty, honesty and integrity are very important from their viewpoint. They can work for long periods of time when necessary in order to achieve goals that are meaningful to them. In addition, they may be perfectionist and forget to recognise theirs and others' efforts. ISTJs usually feel uncomfortable expressing their feelings; as a result they express affection through actions.

As can be observed from Table 5.1, 42.9% of the participants at ITESM-CCM have a *conqueror play style,* whereas 34.3% of the participants have a *manager play style*, according to the DGD1 model. This accounts for approximately 77% of the population. As was discussed in Chapter 3, the *conqueror style* is strongly associated with challenge and enjoys achieving success over adversity. As a result, these individuals possess high tolerance for frustration and have a tendency to finish game tasks. They like to achieve mastery of skills and to perform strategic thinking. The *manager style* does not pursue mastery. They think that if the goal is achieved in any way, it means that they have acquired the necessary skills. There is a high probability that they will not finish the games that they start. They are highly competent dealing with several factors in parallel. They combine tactical competence with strategic thinking.

From Table 5.2 it can be observed that 61.1% of the participants at Trinity College Dublin correspond to the *conqueror play style* of the DGD1 model, while 33.3% of the participants correspond to the *participant play style*. *Participants* have a preference for playing games with other people and like games where they are in charge and that encourage them to be an active part of them, e.g. creating the game story, interacting with game characters, paying attention to detail or being part of a community of players. Specifically, *participants* have a preference for people who are physically present and contribute to the team's emotional state. They do not like competition, but they would compete if the team demands it. They prefer cooperation in order to solve a specific problem or perform a specific task.

Students at both institutions also reported what they believed were the most difficult topics of introductory physics. As is seen in Figure 5.3, the most challenging topics of an introductory Physics course at ITESM-CCM are Kinematics and Dynamics, Vectors in three dimensions, Parabolic movement, Friction and Collision and Momentum. The other topics were mentioned with less frequency or were not mentioned at all. From Figure 5.4 it can be observed that the most challenging topics of an Introductory Physics course at Trinity College Dublin are also Kinematics and Dynamics and Vectors in three dimensions. In addition, they also highlighted the topic, Wave motion. Other topics were mentioned with less frequency or were not mentioned at all.

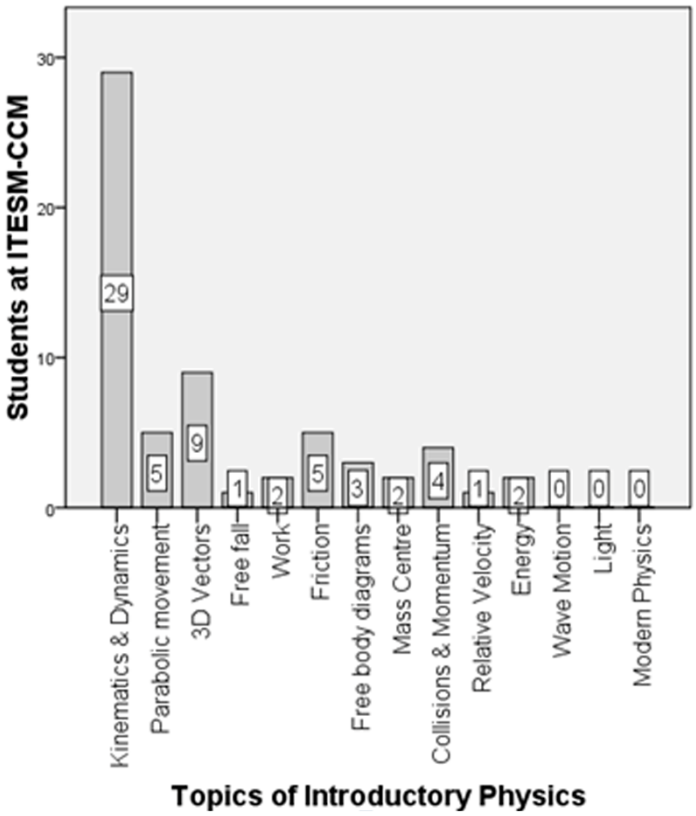


Figure 5.3 ITESM-CCM reports on challenging topics of introductory physics

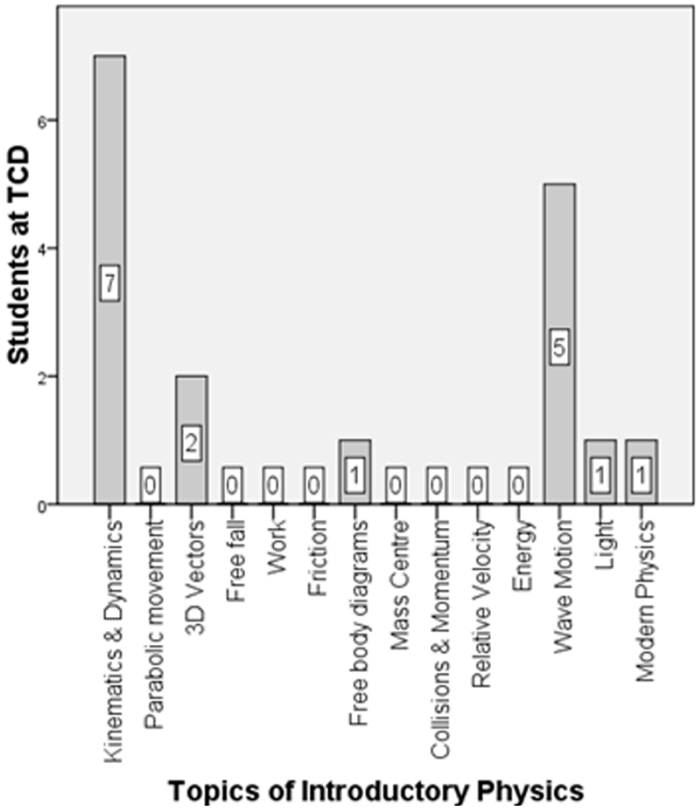


Figure 5.4 TCD reports on challenging topics of introductory physics

In this survey, the hardware equipment available to students at the Universities was assessed. Approximately 49% of the participants perceived the response of the computer equipment at ITESM-CCM as moderate (see Figure 5.5). However, 40% of students rated the equipment as quickly responding. 63% of students reported that the equipment includes microphones and speakers, whereas 91% reported that the computer equipment does not include a camera. When examining the availability of computer equipment at home or self-owned, 34 out of 35 students reported having a desktop computer or laptop. Half of the participants said that they have a desktop computer. 30 students reported having a laptop.

From Figure 5.6 it can be seen that almost 78% of students at Trinity College perceived the computer equipment response as quick. However, this same percentage of students reported that the computer equipment does not have microphones. 56% of students reported that the computer equipment does not include speakers and 89% of students reported that the computer equipment also does not include cameras (see Appendix A). About 74% of students at ITESM-CCM rated that their self-owned equipment responds quickly, whilst almost 24% rated that their self-owned equipment responds moderately (see Figure 5.7). Eighty-five percent of students reported having computer equipment including microphones. However, 91% of students reported that the computer equipment owned included speakers and 74% of the equipment self-owned comprises a camera.

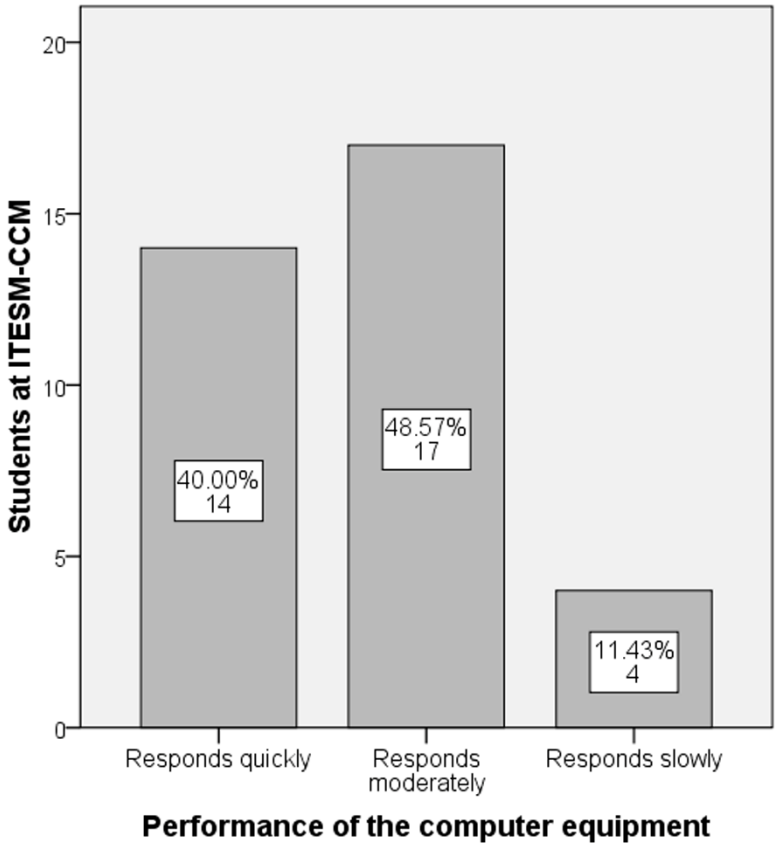


Figure 5.5 Student rating of performance of ITESM-CCM computer equipment

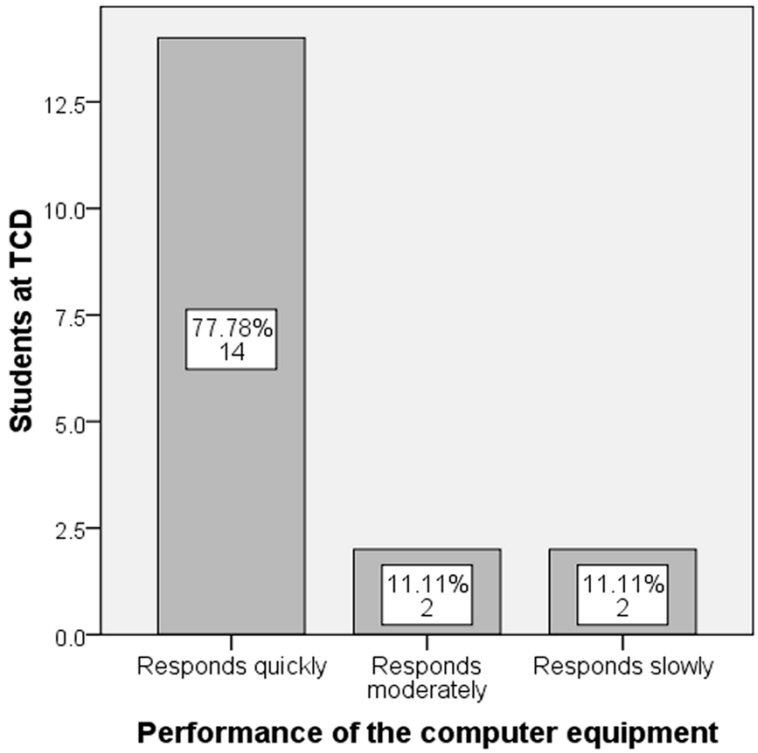


Figure 5.6 Student rating of performance of TCD computer equipment

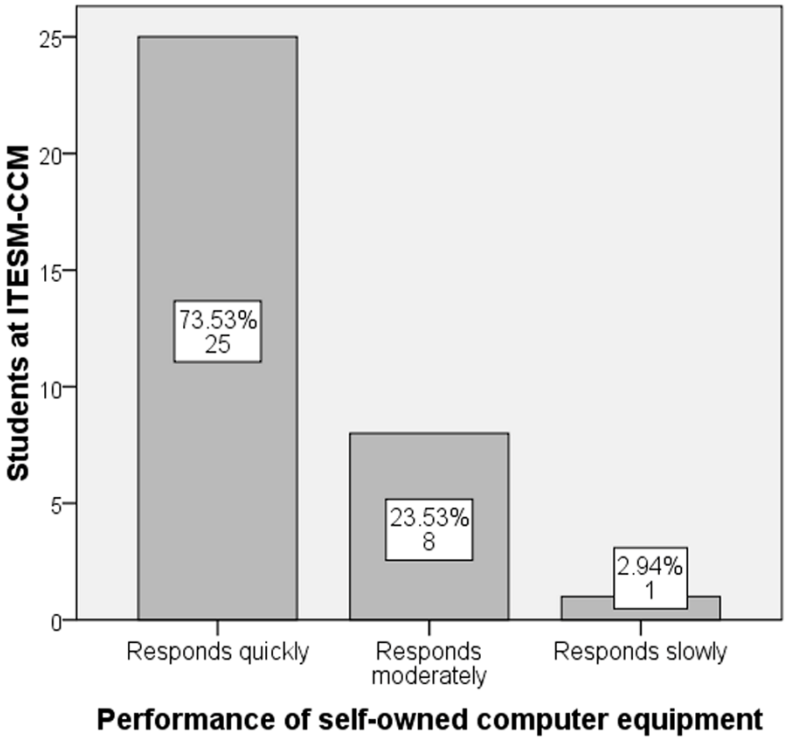


Figure 5.7 ITESM-CCM student rating of self-owned computer equipment performance

Students at Trinity College Dublin reported that 17 out of 18 students have equipment available at their homes. From these 17 students, 10 reported owning a desktop computer and 12 reported owning a laptop. Approximately 65% of students reported perceiving a quick response from their computer equipment (see Figure 5.8). Nobody reported perceiving a slow response from their equipment. 76% of the students reported that their self-owned equipment includes speakers and 53% reported that their self-owned equipment includes a camera (see Appendix A). Students in both institutions reported perceiving that their self-owned computer equipment responds more appropriately and possesses more hardware equipment than university computer equipment.

Additionally, this survey analysed student familiarity and level of confidence with computer and Internet technology and experience playing video games. Students at ITESM-CCM started to use computers when they were approximately 8 years old on average and 30 students out of 35 reported feeling “very confident” or “confident” using them. In addition 31 students out of 35 reported feeling “very confident” or “confident” using the Internet (see Figure 5.9). Almost 29% of students reported that they “always” or “usually” play video games, whilst 34% are occasional players and about 37% hardly ever or never play video games.

It was observed that 80% of students at TCD reported feeling “very confident” or “confident” using computers. Aproximately 94% of students reported feeling “very confident” or “confident” using the Internet (see Figure 5.10). In addition, students started to use computers when they were on average approximately between 10 and 11 years old. 44% of stu-

dents reported playing games “occasionally”, 22% reported playing games “hardly ever” and 33% reported playing games “always” or “usually”. Hence, in comparison with students at ITESM-CCM, there are more students playing video games occasionally (44% vs. 34%) and there were no students who reported never playing games. Similarly, of all participants, about 30% of students play games “always” or “usually”, see Appendix A.

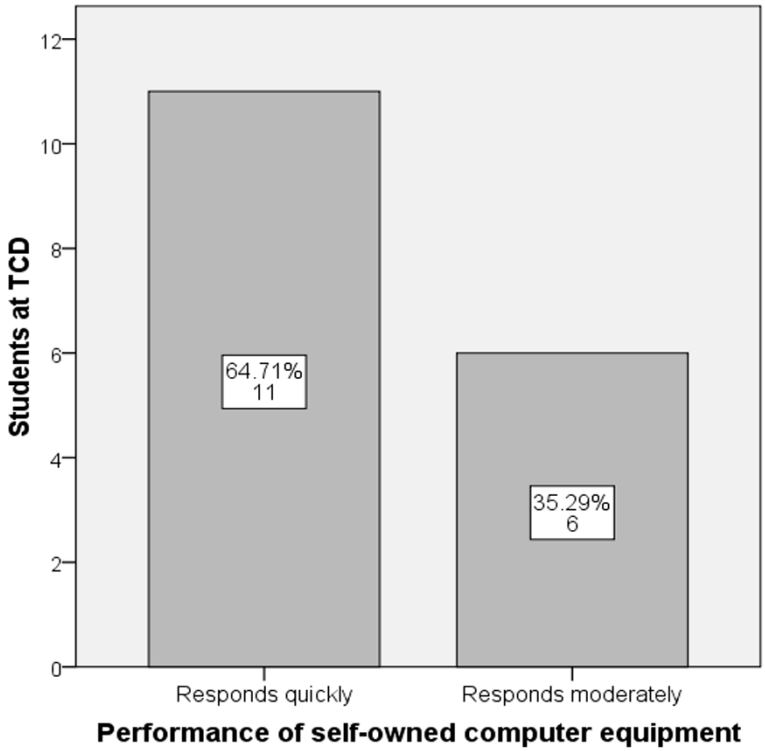


Figure 5.8 TCD student rating of self-owned computer equipment performance

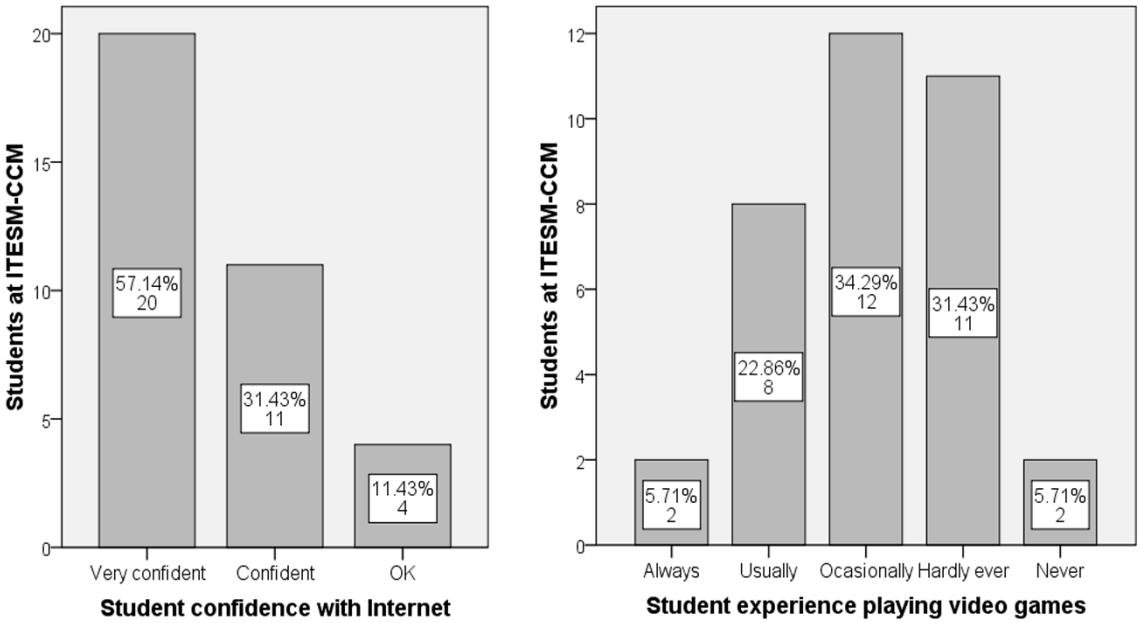


Figure 5.9 ITESM-CCM student experience with Internet and playing video games

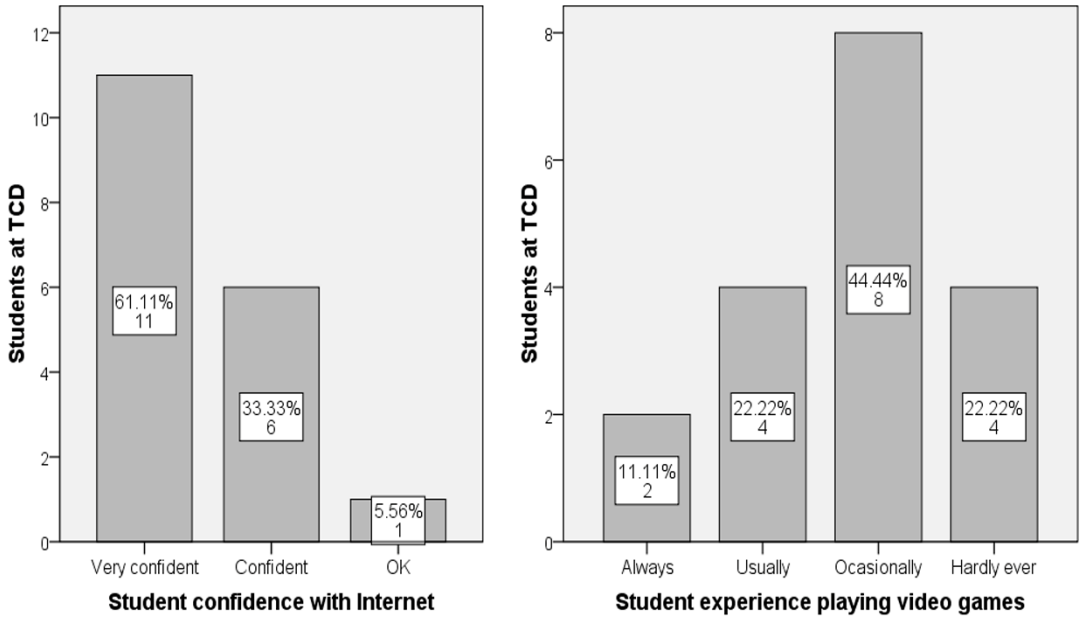


Figure 5.10 TCD student confidence with Internet and experience playing video games

Students were also asked to order the features of an educational computer game according to its influence in enabling them to feel engaged and to facilitate their learning, where 1 is the “most important” and 8 is the “least important”. The average value of student rated importance for each game feature in maintaining engaged and facilitating learning was obtained. According to students at ITESM-CCM, the “goals and objectives” of a serious game are the most important aspect in maintaining their engagement and facilitating learning. Next attention is focused on the “player’s interaction” and the type of “challenges, puzzles or problems”. Students considered that “outcomes and feedback” are the least important game features. Students at Trinity College Dublin agreed with students at ITESM-CCM, when they reported that the type of “puzzles, problems or challenges” and “goals and objectives” are some of the most meaningful elements. However, they also reported as one of these elements, the “player roles” and signalled “game rules” as the least important element. It is possible that students for this question lend more importance to keeping engaged rather than facilitating learning (see Appendix A).

Students were asked to provide examples of features that helped them to feel engaged whilst playing video games. At ITESM-CCM, one student reported that he played a video game involving cars and the interaction and story played a key role in maintaining his engagement. Another student said that video games have helped him to develop skills and to think more quickly and solve puzzles faster. These skills have helped him to solve some real-life problems. Another student said that from her perspective games have to be challenging to test your mental abilities and prompt other ways of solving a problem. Another student reported that undertaking problems as a series of goals and objectives makes the task easier.

Some students signalled that there are some games on Facebook that involve acquiring knowledge and so provided a learning experience. One student reported that he enhanced his learning about World War II through playing the video game “Medal of Honor Frontline”. One student reported the plot of the movie “Mission to Mars” as engaging and from which an educational game could be derived. Some students reported that even though games are challenging, they believe that they do not have to be overly complex in order to experience goal achievement. At Trinity College Dublin, one student reported that the game “Sims”, since it appears life-like during activity performance and for quality of the virtual environment. Another student said that he felt incredibly engaged when playing “Team Fortress 2”, due to the game goals and player roles. Other students suggested games with challenges and goals that are morally relevant, games that encourage planning and thinking on a strategy, challenging puzzles, and games that provide hints and feedback.

Students were also asked to order the probable output/feedback elements of video games according to their perceived importance in influencing their engagement, where 1 is the “most important” and 8 is the “least important”. The average for student influence rating for each element is shown in Tables 5.3 and 5.4. At ITESM-CCM, students believe that “believability” is highly important (see Table 5.3). Then they focus on the “entertainment” and “challenges” provided. However, it was observed that students feel that “emotional responses” or “emotional behaviour”, “colours” and “sounds” are considered the least important elements for engagement.

At Trinity College Dublin, students perceive that “entertainment” or “sense of humour” is very important to maintaining engagement (see Table 5.4). Then students consider that “graphics”, “provision of challenges” and “emotional responses” or “behaviour” also have significant influence over student engagement. However, “sounds”, “colours” and “showing empathy” are considered less important by these students in influencing their level of engagement. The differences between student ratings in both institutions may be due to the cultural difference and the diversity of personality types.

		Colours	*Graphics*	*Sounds*	*Emotional responses or behaviour*	*Empathy*	*Believability*	*Entertainment or sense of humour*	*Provision of challenges*
N	*Valid*	35	35	35	35	35	35	35	35
	Missing	0	0	0	0	0	0	0	0
Mean		5.69	4.17	5.34	5.11	4.40	3.20	4.00	4.09
Std. Deviation		2.259	2.491	1.939	1.510	2.291	2.260	2.339	2.280

Table 5.3 ITESM-CCM student influence ratings on output/feedback elements

		Colours	Graphics	Sounds	Emotional responses or behaviour	Empathy	Believability	Entertainment or sense of humour	Provision of challenges
N	Valid	18	18	18	18	18	18	18	18
	Missing	0	0	0	0	0	0	0	0
Mean		5.61	3.50	5.11	4.44	5.78	4.94	2.89	3.72
Std. Deviation		2.173	1.886	1.811	2.148	1.768	2.667	1.875	2.539

Table 5.4 TCD student influence ratings on output/feedback elements

Furthermore, students were asked to order different elements and strategies that can be used in video and serious games to encourage learning, where 1 is the “most important” and 8 is the “least important”. The average value for each element or strategy was calculated, (see Tables 5.5 and 5.6). “Introduction of new strategies”, “clues” and “incremental increase in the level of difficulty” are considered the key strategies by students at ITESM-CCM. The reported least popular strategies are having “interactive dialogues with animated characters”, “payoffs” and “enquiries”. “Explanations”, “introduction of new strategies” and “embody concepts to be learned” are considered important strategies by students at Trinity College Dublin. They reported that less important strategies are having “interactive dialogues with animated characters”, “payoffs” and “enquiries”, in concordance with reports from students at ITESM-CCM.

		Increase the level of difficulty	Introduction of new strategies	Pay-offs	Embody concepts to be learned	Enquiries	Explanations	Clues	Interactive dialogues with animated character
N	Valid	35	35	35	35	35	35	35	35
	Missing	0	0	0	0	0	0	0	0
Mean		4.03	3.86	5.20	4.60	5.14	4.14	3.89	5.31
Std. Deviation		2.760	2.328	2.139	2.032	2.088	2.144	1.937	2.553

Table 5.5 ITESM-CCM student ratings for influence of elements and strategies

		Increase the level of difficulty	Introduction of new strategies	Payoffs	Embody concepts to be learned	Enquiries	Explanations	Clues	Interactive dialogues with animated character
N	Valid	18	18	18	18	18	18	18	18
	Missing	0	0	0	0	0	0	0	0
Mean		4.28	3.17	5.67	3.22	4.83	2.72	4.56	7.56
Std. Deviation		1.708	2.007	1.940	1.865	1.823	1.602	2.093	1.247

Table 5.6 TCD student ratings for influence of elements and strategies

To know whether specific emotions can be associated with colours, students were asked to relate colours to eight emotions: neutral, happy, sad, disgust, fear, surprise and frustration. 37% of students at ITESM-CCM associate neutral emotion with gray colours, whilst 23% of students associated neutral emotion with blue colours. Happiness is associated frequently with various colours, with 23% of students associating it with orange, whilst 20% of students associate it with yellow, 17% of students associate it with green and 17% of students associate it with blue. Also sadness is related to various colours, 29% of students associate it with gray, 17% of students associate it with blue, 14% of students associate it with black and 14% of students with brown. Disgust is associated with green by 23% of students, 14% of students associate it with orange and 14% of students with brown. Fear is associated with black by 34% of students and is associated with purple by 17% of students. Surprise is associated with yellow by 34% students and is associated with green by 17% of students.

Anger was the emotion on which most agreed - 60% of students associated anger with the colour red (see Figure 5.11). In the case of frustration, we asked students to suggest the colour. Whilst reporting on frustration (see Figure 5.11), it can be observed that students not only reported different colours from those proposed for the other emotions, they also suggested different tones, e.g. green, olive green and oxford green. 14% of students associated frustration with black, 11% with purple and 11% with red.

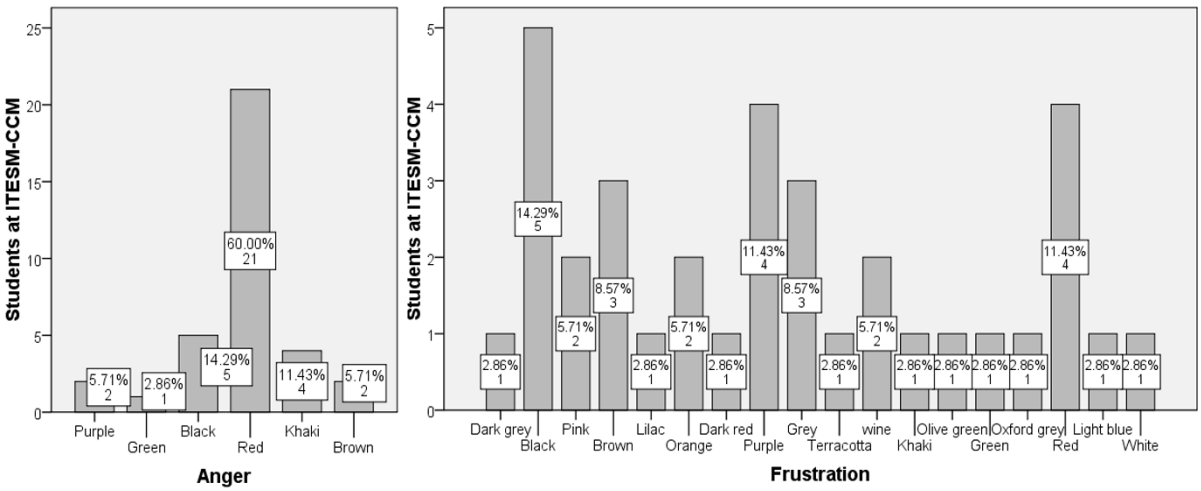


Figure 5.11 ITESM-CCM student associations between colours and emotions

At Trinity College Dublin, approximately 30% of students associated neutral emotions with grey colours, whilst 22% of students associated neutral emotions with green colours. 67% of students associated happiness with yellow, but 22% of students associated it with blue. 39% of students associated sadness with gray, whilst 33% of students associated it with blue. 50% of students associated disgust with brown. 39% of students associated fear with black.

33% of students associated surprise with orange and 28% of students associated it with the colour purple.

Approximately 70% of students at Trinity College Dublin associated anger with red, which is in concordance with results reported by students from ITESM-CCM and it is the emotion with the highest agreement factor, (see Figure 5.12). Also, we asked students to suggest a colour for frustration. They came up with different colours and tones, e.g. 22% of students associated frustration with the colour red, whilst 22% of students associated it with the colour orange (see Appendix A).

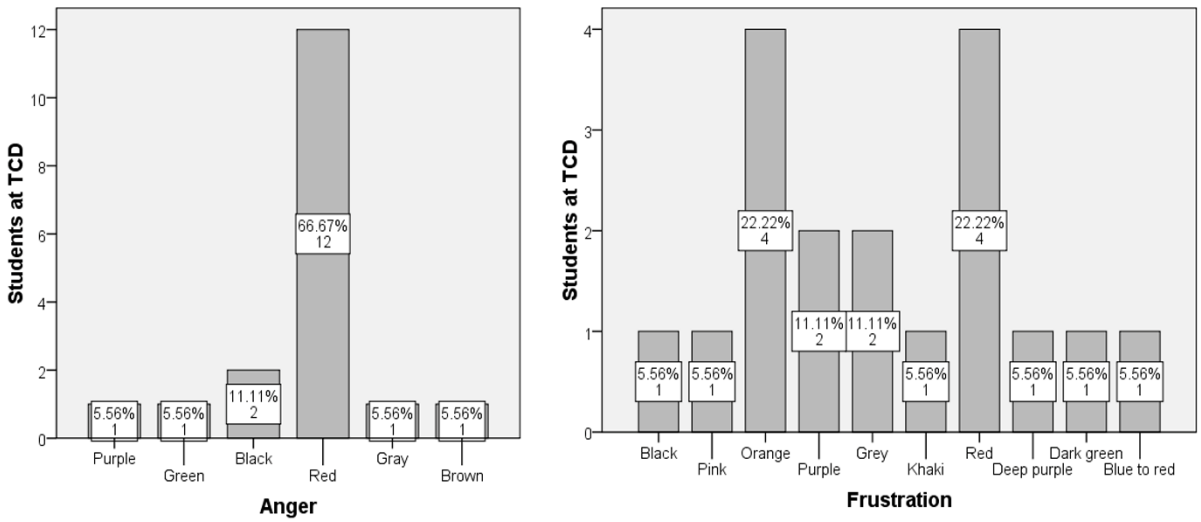


Figure 5.12 TCD student associations between colours and emotions

From these results, it can be concluded that associating emotion with colours is complex, since there is no clear consensus amongst students and also the tone and intensity of colours influences this relation. As was discussed earlier in Chapter 3 innate and cultural meanings are related to colours in addition to students' own meanings based on personal experiences. This may be the key reason why there is no absolute common agreement on associations between colours and emotions. Hence, using colours for conveying an affective message can sometimes lead to ambiguity or misunderstanding.

5.1.2 Lecturer user requirements questionnaire results

Three lecturers from ITESM-CCM and one lecturer from Trinity College Dublin participated in this survey. The lecturers at ITESM-CCM reported on the module Physics I and the lecturer at Trinity College Dublin reported on the module Physics of Motion. Also, the lecturers at ITESM-CCM had groups of 16, 21 and 50 students, whilst the lecturer at Trinity College Dublin had a group of 70 students. In addition, lecturers at ITESM-CCM reported that between 20% or 30% of students fail the module, whilst the lecturer at Trinity College Dublin indicated that only 15% of students fail the module. Also in this survey, we asked lecturers in both Uni-

versities to state the reasons for students failing and the results show concordance. These reasons are:

1. Student lack of motivation, commitment, effort, study and perseverance for overcoming difficult and adverse situations.
2. Inappropriate visualisation of physics phenomena, due to the lack of analytic and structured reasoning of physics case studies.
3. Students have not yet acquired a high proficiency in problem solving skills.
4. Physics concepts are difficult to grasp for some students.
5. Unsuitable elementary or basic Mathematical skills.

Also, we asked two lecturers in both institutions which topics in the module have been identified as the most complex or the ones where students struggle the most. Lecturers at ITESM-CCM said that students struggle specifically with topics encountered for the first time, which comprises the last part of the course or involve applying a number of previous topics taught in this module, such as *the relationship between Conservation Forces and Potential Energy*, and *angular and linear momentum*. The latter topic is specifically complex for students because it involves *Vectors* and this challenges the students. Also lecturers at ITESM-CCM identify *Dynamics of rigid bodies* as a complex topic, since it involves the application of *Newton's laws*, both in *translational and rotational* forms. The lecturer at Trinity College Dublin reported that the topic of *linear and angular momentum conservation* is challenging for students, since it is less intuitive for them. This statement is in concordance with the lecturers at ITESM-CCM.

We also enquired about the fundamental topics of this course. *Energy Conservation* is considered a key topic, since it involves applying many of the topics previously taught in the course and fundamental theorems. *Dynamics*, specifically *Newton's laws* since these comprise the foundation of the course. *Angular and linear momentum* theorems and *Kinematics,* because they describe *movement,* are challenging. These topics are considered fundamental because knowledge of them is required to understand many of the *Mechanics* topics and hence provide the basis for further physics study.

The methods employed to support learning at ITESM-CCM are oral presentations from lecturers, team work in which students must learn through solving specific problems, sitting monthly exams and one final exam; assignments with module questions that students must prepare using the textbook by Serway and Jewett (2004); exemplification of the procedure applied to solve physics problems in classes and workshops; online simulations, questionnaires and links; and resources for learning through mobile devices, e.g. Blackberry smartphone. At Trinity College Dublin, lecturers also use oral presentations, lecture notes and

online assignments. In both institutions, lecturers consider that these methods are sufficient for 80% of students to learn the main goals, objectives, skills and abilities required by the module. In addition, two out of three lecturers at ITESM-CCM believe that allocated module time is sufficient to cover the goals and skills required by the module and the lecturer at Trinity College Dublin also reported that the time available was sufficient to obtain the same outcomes. One lecturer at ITESM-CCM suggested that an additional lecture session each week can make a difference and more self-study resources, such as tutorials and simulations, may help in aiding students to grasp concepts and understand procedures. Lecturers suggested the following ideas for a game where students can learn physics concepts:

1. A car driving over a highway with loops or a rollercoaster, since a number of laws of physics can be associated with it.
2. A parachutist jumping from a platform or a person bungee jumping.
3. A game where the concepts behind the application of Newton's laws for both particles and rigid bodies are taught.
4. A game where the concepts behind energy and linear or angular momentum conservation are taught.
5. A strategy game based on decisions about physics.
6. A mystery or detective game where a crime will be solved through interrogating suspects, since experiments are the only way that Physicists have to interrogate nature.

Lecturers also reported that stories in games can be developed around activities that students perform in their everyday lives. However, the lecturer at Trinity College Dublin indicated that he is rather sceptical about using games for teaching purposes.

5.1.3 Key findings from user requirements analysis questionnaires

The key findings from student and lecturer user requirements analysis online questionnaires are:

1. More than half of the students in both institutions, Trinity College Dublin and ITESM-CCM, reported feeling very comfortable "learning by doing", i.e. tactile or kinaesthetic learning.
2. The majority of students in both institutions, e.g. ~43% students at ITESM-CCM and ~60% students at Trinity College Dublin match with the DGD1 model *conqueror play style* (see Table 5.1 and 5.2).

3. *Kinematics* and *Dynamics*, which involve the application of *Newton's laws for particles and rigid bodies; and Vectors in Three-dimensions (3D)* were identified by students in both institutions as the most complex topics of an introductory physics course at undergraduate level. Lecturers also suggest that Dynamics is complex for students, because it involves the application of *Newton's laws in translational and linear forms and angular and linear momentum* are also difficult for students, since they involve *Vectors in 3D*.
4. Students at ITESM-CCM started to use computers when they were on average 8 years old, whilst students at Trinity College Dublin started to use computers when they were on average between 10 and 11 years old. As a result, ~80% of students in both institutions reported feeling very confident using computers and ~90% of students at both institutions reported feeling very confident using the internet.
5. ~30% of students at both institutions often play games and another 30% or 40% of students at both institutions play games occasionally.
6. Students at both institutions consider that *goals and objectives* and *puzzles, problems or challenges* are key elements of educational games that influence their engagement and facilitate learning.
7. Output/feedback elements of educational games that are considered key to influencing student engagement in both institutions are entertainment or sense of humour and the provision of challenge.
8. Elements and actions used in video and serious games that are considered key to influence student learning at Trinity College Dublin are *explanations*, *the introduction of new strategies* and *embody concepts* whilst students at ITESM-CCM are also in concordance with the *introduction of new strategies*, in addition to the use of clues.
9. On the association of colours with emotions student questionnaire results from both institutions, show that sometimes using colours to convey a message in a video game can be ambiguous and uncertain, as some colours of different tones and shades can be associated with many emotions. However, from our questionnaire results, 60% of students suggest that red relates to *"anger"*, which also suggests a dependency on the emotion that wants to be conveyed.
10. One out of the three lecturers at ITESM-CCM believes that the time available is not sufficient to enable students to grasp concepts and achieve necessary skills. Giving an additional lecture session each week and providing more self-study resources, such as tutorials and online simulations was suggested.
11. Lecturers at ITESM-CCM use diverse resources, such as online and mobile technologies, to support student self-study. Therefore, it can be inferred, that they are

open to the use of modern methods of teaching, whilst lecturers at Trinity College Dublin have a preference for more traditional methods.

5.2 Storytelling and fantasy within PlayPhysics

In accordance with the recommendations of lecturers and students at Trinity College Dublin and ITESM-CCM, PlayPhysics focuses on teaching the topics of Kinematics and Dynamics, specifically Newton's laws for particles and rigid bodies and vectors in 3D, which are considered the most challenging topics at the first year undergraduate level. Given that more than half of the students at ITESM-CCM feel *very comfortable* with learning by doing, it would be desirable to fulfil all student expectations related to the game genre. However, it is time-consuming and complex to create a game that suits all. As a result, it was decided to create a game that mainly fulfils the expectation of the *conqueror play* style of the DGD1 model, since ~40% of students match this category.

To target the 'conqueror' play style and in accordance with Control-value theory, PlayPhysics is comprised of game challenges, where each challenge corresponds to a problem solving activity that is related to the achievement of specific learning goals, i.e. outcomes. From the suggestions provided by students to derive the game fantasy or story (see Appendix A), it was observed that students like a combination of real facts with sci-fi, space and action adventures. For example, a student mentioned that a story like the one developed in the movie "Mission to Mars" would be interesting. As a result, PlayPhysics is designed as a role-playing game (RPG) where the student is an astronaut in a mission and as part of this mission must overcome tasks that involve the understanding and application of physics concepts.

In PlayPhysics, the student is an astronaut with the mission of saving his/her mentor, Captain Foster, who is trapped on space station Athena. Captain Foster is injured, and was unable to escape Athena with the rest of the crew. The super computer, VNUS-27-81 (from now on, VNUS) attacked Athena's crew. The source of the problem was a computer virus. The mission starts when the student is ready to be launched from Earth to travel to Athena. Figure 5.13 shows an image from PlayPhysics' story plot.

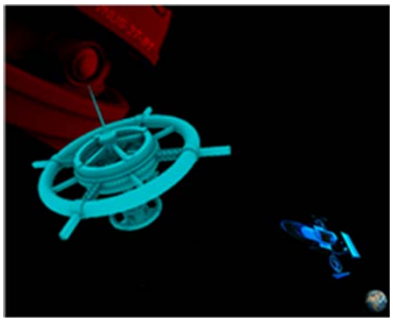


Figure 5.13 PlayPhysics story plot image

Athena is a space station, with a donut shape and rotating at constant acceleration, located between Mars and Jupiter. In order to complete the mission, students must overcome game challenges that involve application of their knowledge and understanding of physics. Each challenge is associated with a single game level (see Appendix B). Here, we are concerned with the design of the first level game challenge as this is sufficient for thesis objectives given in Chapters 1 and 4. Two player characters, male and female astronauts, were also designed (see Appendix C).

5.3 Specifying the physics domain

A domain expert in Astrophysics, Professor Luis Neri from ITESM-CCM (Mexico), participated in specification of the PlayPhysics game scenarios, which are designed so that students can interact from a first person perspective. In the first game challenge, the Alpha Centauri spaceship, having been launched from the Earth, is heading at constant speed towards the Athena space station. The main goal of this case study is that the student chooses and sets appropriate values for physics variables to stop Alpha Centauri at Athena's rotational axis. This goal has to fulfil conditions such as determining a position that facilitates docking and entering Athena before Alpha Centuri's fuel is exhausted.

The case study is presented by a learning companion, referred from here on as *M8-robot*. M8-robot can ask questions of students and provide collaborative learning strategies that assist students to reflect on their own learning and conduct independent learning in order to apply it to other contexts or situations (Woolf 2009). Simultaneously, it is an approach that enables students to participate in social interaction to achieve precise learning outcomes. As a result, *M8* will provide direction through hints when required, ask students to report frequently their emotional state, mirror student affective behaviour or provide a motivating message.

The teaching concepts of PlayPhysics' first game challenge are related to the topic of one-dimensional rectilinear motion, which is one of the core topics of an introductory physics course at undergraduate level. In this specific case, to achieve the goal of this game scenario, constant deceleration must be applied. The initial distance (D) from Alpha Centauri to Athena and the remaining time (T) to exhaust the fuel are constraint variables. These variables are set initial random values by PlayPhysics. However, they are assigned within specific value ranges $D \in [17, 50]\ km$ and $T \in [80, 120]\ s$. These value ranges were defined through an elaborate and precise analysis in order to maintain a non-trivial solution. As part of the game-play, students have to define the values of the exploration variables in order to understand the effects of their decisions and identify the appropriate combination that will achieve the goal of the game scenario. Students must concentrate on the direction of Alpha Centauri's acceleration - towards (←) or away (→) from Athena, the magnitude of Alpha

Centauri's acceleration (a) and its initial velocity (v_i). The magnitude of Alpha Centauri's acceleration has to be within a range $a \in [0, 100]$ m/s^2 and its initial velocity has to be within a range $v_i \in [1000, 2000]$ m/s.

By analysing the case study, students need to realise that they have to first focus on selecting the direction of acceleration. In this case they must understand that choosing a direction towards (←) Athena, will only make Alpha Centauri accelerate until the fuel is exhausted. As a result, Alpha Centauri will continue moving in interplanetary space forever. Hence, the correct response involves setting the Alpha Centauri's acceleration direction away from Athena (→). If students make the error of setting incorrectly Alpha Centauri's direction of acceleration, immediate and implicit feedback is provided by displaying the main character with a face expressing desperation, saying 'I am lost...' and showing the main character thinking of the infinity symbol (∞). Additionally, if the student asks M8 to assist him/her in clarifying their misconception, M8 will answer: 'Oops...Alpha Centauri did not stop at Athena's axis. Its speed continues to increase'. It is important to signal that if the student does not ask M8 to clarify the misconception, M8 will not offer the information voluntarily. The main purpose of this is to allow more independent students to come up with the solution without any help.

After the direction of Alpha Centauri's acceleration is correctly selected, student attention is centred on determining the value corresponding to the magnitude of Alpha Centauri's deceleration (a) and initial velocity (v_i) considering the previously defined value ranges. In the case of the deceleration magnitude, this has to be set with a value no larger than $4g$, i.e. $a \leq 4g$, where g represents the gravitational acceleration of the Earth $g = 9.81$ m/s^2. Since students may feel persuaded to select a large value for Alpha Centauri's deceleration magnitude in order to make it stop as soon as possible and not depleting all available fuel, they must realise that humans cannot tolerate accelerations larger than $4g$. If the student makes this error and asks M8 for clarification, M8 explains to the student that, 'It seems the magnitude of the acceleration is too large (more than $4g$) and as a result, the astronaut blacked out'. Therefore, the student has to select a lower value for the magnitude of the acceleration. In this specific case, the immediate and implicit feedback provided is to display the player character in a purple colour with a face expression of death and thinking, 'Too much acceleration'.

Alternatively, if the student selects a small value for the magnitude of the acceleration (a), the student is very likely to go beyond the time limit (T), which is a time constraint for completing the game challenge. PlayPhysics must calculate Alpha Centauri's breaking distance (d_s) and the time required to stop (t_s), applying Equations 5.1 and 5.2, in order to evaluate the effectiveness of student decisions when choosing the values for a and v_i (see Appendix D).

$$d_s = \frac{v_i^2}{2a} \qquad \text{Eq. 5.1}$$

$$t_s = \frac{v_i}{a} \qquad \text{Eq. 5.2}$$

To start the evaluation of the given solution, firstly, PlayPhysics compares t_s against T. If t_s is larger than T, i.e. $t_s > T$, means that the student chose values that caused the fuel to be exhausted before Alpha Centauri arrived at Athena. In this case, PlayPhysics awards the student a low grade and provides immediate feedback by displaying the main character looking through binoculars and saying, 'It is too far'. Then, if the student asks M8 for clarification about the misconception, he says: 'Alpha Centauri stopped too short from Athena. Choose other values for v_i or a.

If t_s is smaller or equal to T, i.e. $t_s \geq T$, PlayPhysics must calculate the relative error of the distance (e_d) (see Equation 5.3) by subtracting from the breaking distance (d_s), the initial given distance (D) from Alpha Centauri to Athena station. The latter divides the result of the subtraction calculating the relative error, which is multiplied by 100 in order to obtain a percentage.

$$e_d = \frac{d_s - D}{D} \times 100 \qquad \text{Eq. 5.3}$$

As student calculations become more accurate, the relative error will become less. Higher scores are given to students who provide suitable solutions corresponding to small values of e_d. The highest score is assigned to solutions in which the absolute value of the relative error $|e_d|$ is lower than a relative error of 2%. However, when $|e_d|$ is larger than a relative error of 10% a low score is assigned to the student solution. Another aspect, which must be considered to evaluate the student solution, is the resultant breaking time t_s associated with each set of randomly assigned constraint variables, T and D. The procedure to evaluate t_s entails calculating the corresponding time interval (Δt) for all possible values of t_s that are consistent with valid values of a and v_i. Cases in which less fuel was consumed, corresponding to smaller values of t_s, are assigned higher scores (see Appendix D).

It can be observed that providing a suitable solution for this game scenario or case study is not trivial, since the following factors must be considered:

1. Selecting the correct direction for applying the acceleration, i.e. deceleration.

2. Not surpassing the maximum acceleration that can be tolerated by humans before passing out, i.e. $40\ m/s^2$.
3. Not exceeding the fuel exhaustion time ($t_s \leq T$) and achieving the lowest relative error e_d in the breaking distance (d_s).
4. Determining the lowest value for the breaking time (t_s).

5.3.1 Analysing student behaviour to assessing knowledge

Events are the effect of student actions, which may be related to student knowledge and understanding, but also to student misconceptions. As a result, interaction events related to student problem solving skills and capabilities must be identified and categorised in order to assess the quality of student decisions. In order to facilitate the assessment of student misconceptions and interaction with the game challenge, a 'Control instruments panel' was designed, where students can explore and set different values and directions for Alpha Centauri's vector quantities: initial velocity (v_i) and acceleration (a), which must satisfy the conditions given in section 5.1.3 in order to complete the game challenge and learning goals.

Hence using the physics domain analysis as a reference and with the help of our domain expert, we defined a marking scheme for quantitatively evaluating student decisions which is represented with nine production rules summarised in Table 5.7. It is important that the time to exhaust the combustible (T) and the distance to Athena (D) are initialised randomly by PlayPhysics, and the constraints of this game challenge. The production rules in Table 5.8 are given in the order of the problem solving procedure as conducted by an expert student. First, the concept of *acceleration* is assessed, afterwards the concept of *time* and finally *distance*. Time and distance involve the concept of *Velocity*.

To assess distance, the time travelled (t_s) by Alpha Centauri must be less than the time to exhaust the combustible (T). The relative error of the distance (e_d) is calculated in order to evaluate how far away Alpha Centauri is from Athena. Also a minimum time $(t_{\min})$ and a maximum time $(t_{\max})$ are calculated in order to define a minimum interval of time (Δt) for accurately assessing the quality of the student's solution. The minimum time is calculated with the maximum value that can be set for initial velocity, i.e. $2000 m/s$. The maximum time is obtained through comparison of the time calculated with the minimum value that can be selected for initial velocity, i.e. $1000 m/s$ and time to exhaust the combustible. The interval of time is obtained from the absolute value of the maximum time less the minimum time.

Rule #	Physics Concept	Misconception	Production rule		Evaluation (mark)
1	Acceleration	Direction of Alpha Centauri's acceleration set incorrectly	*if(a>0)* Accelerating instead of decelerating		0*
2		Magnitude of the acceleration is set incorrectly	*if(|a|>4g)* Exceeds the acceleration that person can tolerate		50
3	Time (Velocity)	Misunderstanding of the relationship between the time that Alpha Centauri travels (t_s) before stopping, its initial velocity and acceleration	*if(t_s>T)* The time that Alpha Centauri travelled exceeded the time to exhaust the combustible		50
4	Distance (Velocity)	Misunderstanding of the relationship between the distance that Alpha Centauri travels (d_s) before stopping, its initial velocity and acceleration	*if(-2≤Round(e_d)≤2)* The relative error of the distance travelled by Alpha Centauri is approximately = 2%	How near or far from the percentage range?	100
				$t_{\min} \le t_s \le \left(t_{\min} + \frac{\Delta t}{3}\right)$	
5				$\left(t_{\min} + \frac{\Delta t}{3}\right) \le t_s \le \left(t_{\min} + \frac{2\Delta t}{3}\right)$	95
6				$\left(t_{\min} + \frac{2\Delta t}{3}\right) \le t_s \le \left(t_{\min} + \Delta t\right)$	90
7			*if(-2>Round(e_d)≥-5)* or *(2<Round(e_d)≤5)* The relative error of the distance is 2% but ≤5%		80
8			if*(-5≤Round(e_d)≤-10)* or *(5≤Round(e_d)≤10)* The relative error of the distance is >5% but ≤10%		70
9			*if(-10>e_d>10)* The relative error of the distance is larger than 10%		60

Table 5.7 Production rules for quantitatively assessing student knowledge

5.4 Requirements analysis for PlayPhysics

This section discusses the requirements for PlayPhysics in terms of testing our hypotheses and satisfying user needs.

5.4.1 Functional and non-functional requirements

From the analysis of user requirements we identified functional and non-functional requirements summarised in Tables 5.8 and 5.9. PlayPhysics, an online game-based learning envi-

* Cases that are strongly penalised by PlayPhysics.

ronment, is being created with the purpose of assisting students undertaking an Introductory Physics module at ITESM-CCM to conduct problem-based learning in an engaging manner, enabling them to explore various physics phenomena. PlayPhysics is not aimed at substituting lecturers and lectures, but endeavours to support lecturer instruction and independent student learning. Hence, it is anticipated that students will enhance their learning about physics by interacting with PlayPhysics.

Lecturers require that the game challenge physics behave as they would in real life scenarios. PlayPhysics will focus on teaching the topics of Kinematics and Dynamics, principally the application of Newton's laws to particles and rigid bodies. Some examples may involve concepts of Vectors in three dimensions, since students consider this topic challenging. Between 60% and 70% of the students participating in our research have experience and familiarity with commercial video games. Educational games are often biased towards the goal of *'teaching subject content'* than towards the goal of *'having fun'*. Given that the students have experience playing commercial video games and not educational games, this may present a challenge, since their expectations may be more biased towards *'having fun'*. Special effort will be placed on identifying learning goals and game challenges that include common game design techniques, since students report that they are important for maintaining engagement. Visual cues will be employed in order to provide direction. Colours will be used in conjunction with images according to Human-Computer Interaction (HCI) guidelines, but they will not be used to convey emotions, since they can have ambiguous emotional meaning for students. Hence, feedback will be conveyed in the game mainly by animations, graphics and sounds.

Lecturers indicated that they prefer that students will interact with the GBL environment online. As a result, PlayPhysics should include functionality that enables students to register, can handle personal details and allow students to securely log in and out. Also, it is required that students advance automatically and at their own pace between activities, e.g. pre-test, game challenges, qualitative questionnaires.

The Head of Department and lecturers of physics at ITESM-CCM anticipate that a System Administrator will handle the registering and maintenance of semesters, subjects, other System Administrators and Heads of Department in PlayPhysics. Lecturers require that PlayPhysics identifies student misconceptions and provides suitable instruction, which depends on student independence. It is assumed that more independent students require less explicit forms of feedback, such as visual, written and auditory responses whilst dependent students will require that their misconceptions are described more clearly and will expect more elaborate hints.

Functional requirements
• Physics must behave as expected in real-life scenarios • Students should be able to interact with PlayPhysics GBL environment online • PlayPhysics should focus on teaching the topics of Kinematics and Dynamics, principally the application of Newton's laws to particles and rigid bodies • PlayPhysics should include functionality that enables students to register, that can handle personal details and that allows students to securely log in and out • PlayPhysics must enable students to advance between activities automatically at their own pace • PlayPhysics must enable the System Administrator to register and maintain semesters, subjects or other System Administrators and Heads of Department • PlayPhysics must identify student misconceptions and provide suitable instruction according to student independence • PlayPhysics should enable students to self-report their emotional state at any time • PlayPhysics must be capable of requesting that students self-report their emotions • PlayPhysics should provide an affective response, i.e. mirroring students self-reported emotional state using NPCs • PlayPhysics must be capable of capturing student physiological data • A version of PlayPhysics that only presents the student with a PowerPoint presentation with the concepts taught by PlayPhysics GBL environment must be implemented • Students should be able to answer a pre-test and post-test to measure their knowledge before and after interacting with PlayPhysics • PlayPhysics should employ clues or hints in order to provide guidance

Table 5.8 PlayPhysics functional requirements

For investigation purposes, PlayPhysics is expected to enable students to voluntarily self-report their emotional state at anytime during interaction and when they are requested to do so. The goal is to collect the necessary student interaction data to complete the definition and evaluation of PlayPhysics' emotional student model. It is also anticipated that PlayPhysics will provide affective feedback in response to student self-reported emotions. This can be achieved with NPCs, which can play the role of learning companions that encourage and mirror student emotional states. The objective is to understand whether students consider emotional responses from NPCs contextually appropriate.

In some cases, PlayPhysics is required to capture student physiological data, such as GSR signals, to determine if the accuracy of the model can be increased through this modality. Hence, PlayPhysics will receive physiological data from a biofeedback device and record it in a database when a student is interacting on-site. In order to assess whether students learn with PlayPhysics, a version of PlayPhysics that only presents the student with a PowerPoint presentation with worked examples and concepts taught by PlayPhysics GBL environment must be implemented. Students who view the presentation are also required to solve the pre-test and post-test.

Non-functional requirements
• Enable students to conduct problem-based learning in an engaging manner • PlayPhysics should comprise intuitive Graphical User Interfaces (GUIs) that enable students to explore physics phenomena and effects of their own actions • Support lecturer instruction and student independent learning • It is expected that PlayPhysics will help students enhance their understanding of physics • PlayPhysics affective responses should be contextually appropriate • PlayPhysics should be comprised of game challenges with clear learning goals

Table 5.9 PlayPhysics non-functional requirements

5.4.2 Use cases

The functional requirements of PlayPhysics can be modelled and visualised with Unified Modelling Language (UML). PlayPhysics use cases corresponding to *Generic User*, *System Administrator*, *Head of Department and Lecturer* and '*Student'* are modelled with UML. PlayPhysics use cases are shown in Figures 5.14, 5.15, 5.20 and 5.21. These are created with Visual Paradigm (Visual Paradigm 2011), a Computer-Aided Software Engineering (CASE) tool. The *Generic User* use case, Figure 5.14, exemplifies the operations that every user must perform in PlayPhysics, e.g. log in and log out. As it can be observed, the other use cases, i.e. system administrator, head of department and student, extend from this Generic User use case.

Figure 5.15 shows the System Administrator use case, who performs activities such as the organisation of Semesters, Subjects or Staff, and also, connect or disconnect a student to a Bluetooth GSR sensor. PlayPhysics manages and monitors the GSR connection through a separate thread. As long as the connection is alive, the thread will record the GSR raw value in the PlayPhysics database. Each student interacts with PlayPhysics at his/her own pace. To avoid students accessing the post-test and pre-test at the same time, since this could encourage students to cheat and bias the results, a Finite State Machine (FSM), which automates this flow of interaction, is implemented. The administrator can manipulate and manually advance the flow of interaction of the student from one stage to another. For example, if the student is working currently with the work tray in the state 'Pre-test' (see Figure 5.16) the administrator has the authority to advance the student to the stage 'Game', 'Post-test' or 'No-access'.

The state 'Pre-test' can be an initial state, as indicated by the arrow '→', and it is where students can observe in their work tray the active web link entitled "Do pre-test" (Figure 5.17) and hence access it. Then the student can move to the 'Game' state or the 'Presentation state' depending on which group to which they belong, namely control (C) or focus (F). In the 'Game' state, students can access the web link "Launch PlayPhysics" (see Figure 5.18) in

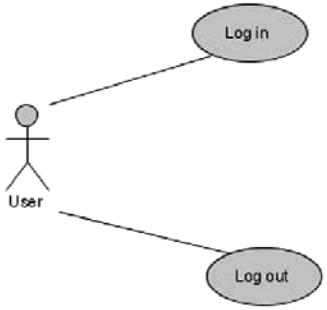


Figure 5.14 PlayPhysics Generic User use case

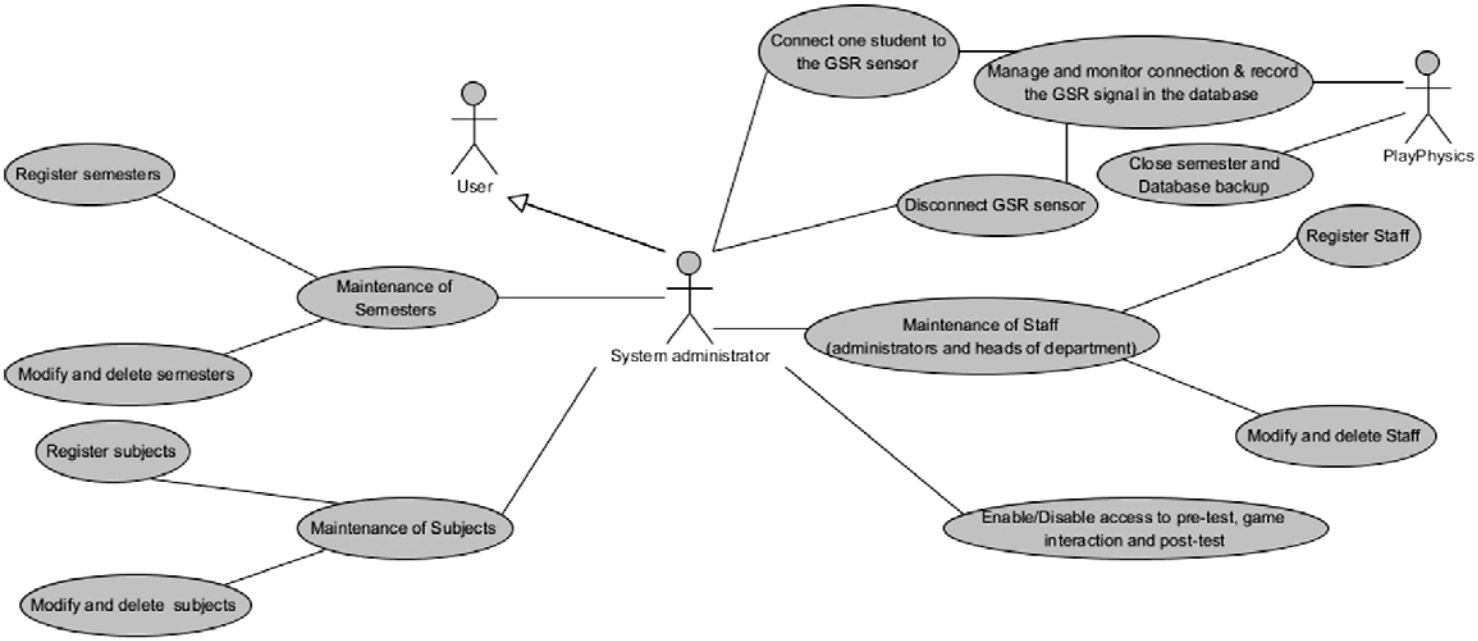


Figure 5.15 PlayPhysics System Administrator use case

order to enter the educational game, whilst 'Presentation' enables access to a power point presentation with problem solutions.

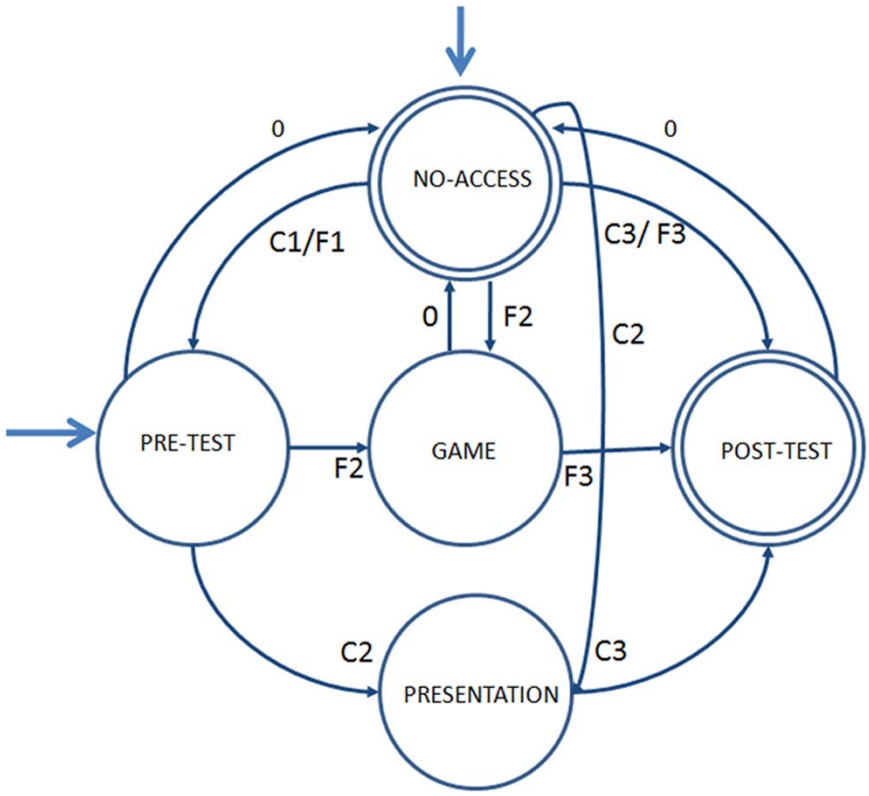


Figure 5.16 Finite State Machine (FSM) of student work tray

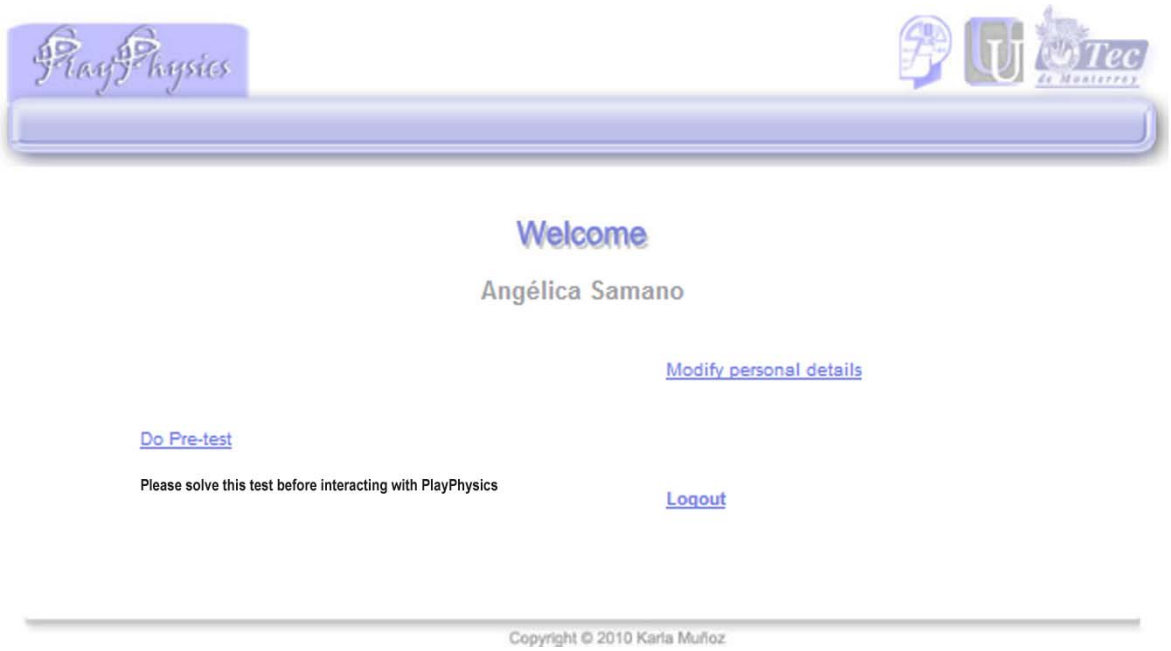


Figure 5.17 Student work tray showing available links in 'Pre-test' state

Once students have started to interact with PlayPhysics and have arrived at the 'Game' stage, they can interact with the game challenge. The work tray then advances to the state 'Post-test' where the corresponding links "Do Post-test" and "Do Qualitative Questionnaire" appear in the work tray and can be accessed when they are playing the game (Figure 5.19). In a similar manner, when students decide they have finished the learning stage, they can advance to the 'Post-test' state. Whenever students feel prepared to take the pre-test, they can progress to it. They have just one opportunity to take the exam. Once students have answered the post-test and qualitative questionnaire, the links are no longer available. There is a fourth state where the student has access only to the links: 'logout' and 'modify personal details'. This state is 'No-access', which is one of the two possible initial states of the work tray Finite State Machine (FSM) and also corresponds to the final state of the work tray.

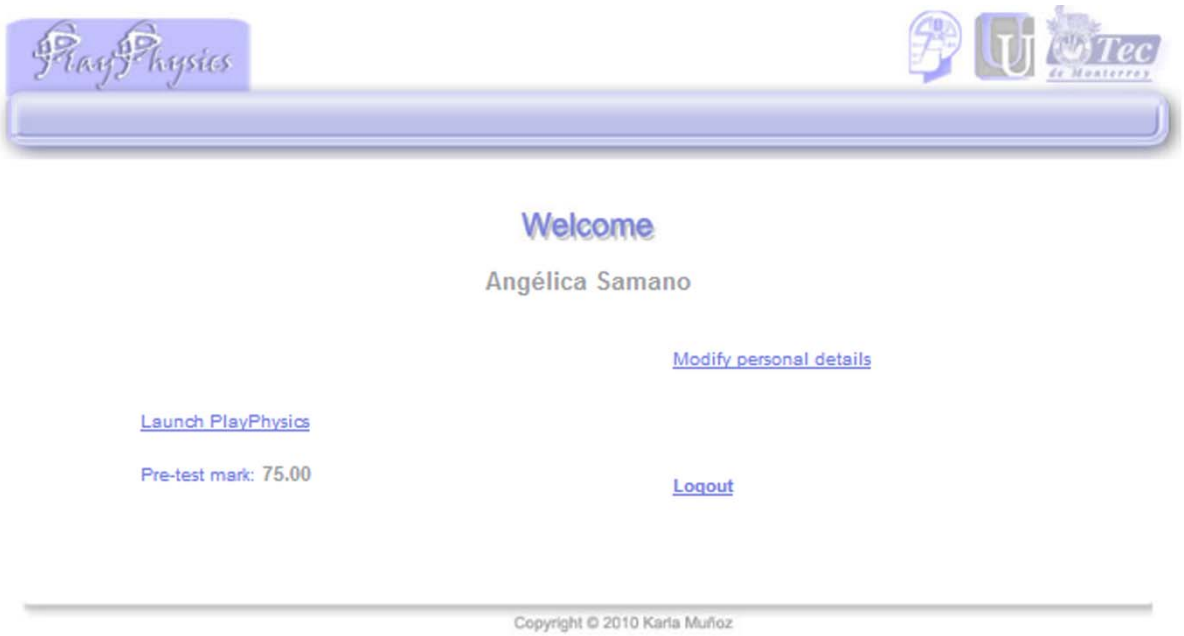


Figure 5.18 Student work tray showing available active links in 'Game' state

The 'C' and the 'F' letters in the transitions, e.g. *C*1 or *F*1, denote the groups in which students are located, i.e. "Control" or "Focus". Students in the control group can advance to the state "Presentation" and access a presentation with physics concepts and some sample exercises, whilst students in the focus group may advance to the state "Game" to access the PlayPhysics educational game. Figure 5.20 shows the Head of Department and the Lecturer use case. The Head of Department and Lecturer share similar functionality, but the former also has the authority to maintain groups and lecturers, i.e. register, delete and modify groups and lecturers. Different functionality is maintained through the use of personalised work trays. The purpose of creating lecturer and head of department user categories is to allow them to review their students' interaction and level of performance.

The *Student* use case is shown in Figure 5.21, which shows functionality and interaction between the Olympia architecture and the student. The student can “register with the system”, “modify personal details and data”, “do a pre-test”, “do a post-test”, “answer a qualitative questionnaire”, “publish his/her game outcomes”, “self-report his/her emotional state and “receive and evaluate the quality of the feedback provided”. On the other hand, Olympia is capable of “examining student interaction and behaviour”, “choose pedagogical, affective or motivational actions” and accordingly “adapt via dynamic interaction module” in order to provide a suitable response.

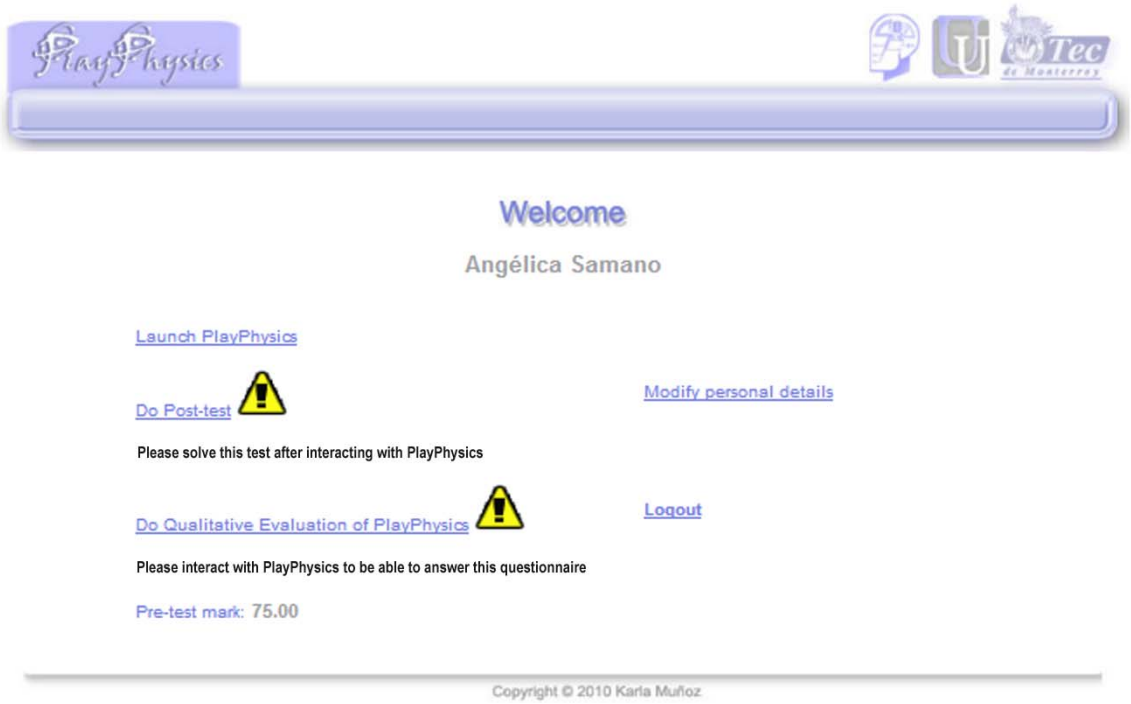


Figure 5.19 Student work tray in `Post-test' state

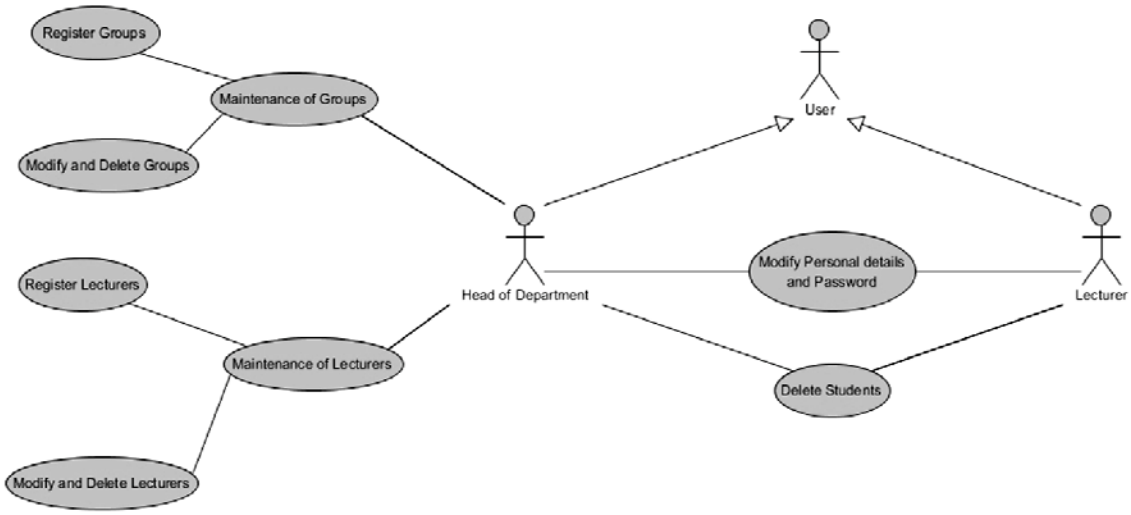


Figure 5.20 PlayPhysics *Head of Department* and *Lecturer* use case

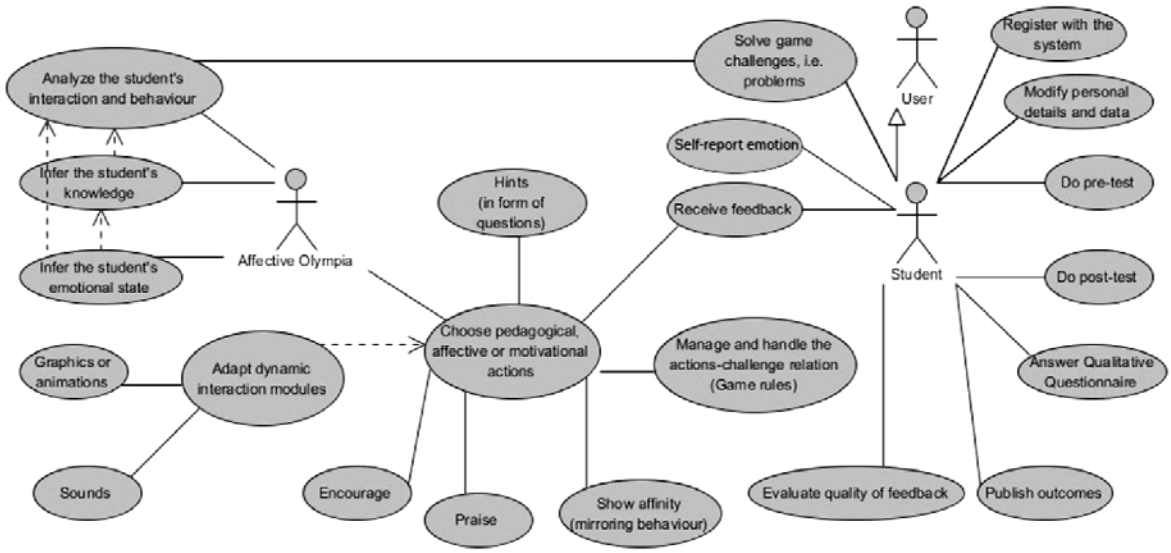


Figure 5.21 PlayPhysics *Student* use case

5.5 PlayPhysics design

This section discusses the detailed design of PlayPhysics, focusing on the emotional student model, software architecture, conceptual representation of data and user interaction.

5.5.1 Emotion recognition: PlayPhysics' emotional student model

PlayPhysics is an application where our proposed emotional student model can be instantiated and evaluated. The application of Binary or Multinomial Logistic Regression (MLRs) and the definition of all the relations in the skeletons of the Bayesian Belief Networks (BBNs) comprising a dynamic sequence of BBNs and corresponding to time frames before, during and after, require acquisition of historic data. With this objective in mind, PlayPhysics has to include functionality that ensures that students' self-report their emotional state frequently and keep track of observable random variables related to student *achievement emotions*. The manner in which our emotional student model is populated for the PlayPhysics application is presented here.

The BBNs corresponding to *prospective outcome, activity* and *retrospective outcome* emotions, which are related to the time frames *before*, *during* and *after* performing a learning activity respectively, are shown in Figures 5.22 to 5.24. In PlayPhysics, the learning activity refers to each game challenge. These PRMs were obtained by substituting the random variables chosen and described in Chapter 4, Section 4.1 into the generic network skeletons described in Chapter 4, Section 4.3. As can be observed in PlayPhysics' Prospective outcome emotions network (Figure 5.22), the student focus on beliefs and attitudes is already apparent before interaction with the game challenge commences. However, given the de-

scription of the topics and activity to be performed, the student must reflect on the probable level of difficulty and his/her confidence in achieving a positive outcome. Hence, we ask students questions related to these random variables during game dialogues, before each game challenge. The game dialogues must encourage students' reflection on their own experience, proficiency and capabilities and simultaneously give a description of the task to be accomplished, goals and challenges involved.

PlayPhysics' activity emotions network (see Figure 5.23) focuses on the current state of interaction variables. However, it must be noted that the acquisition of these variables requires PlayPhysics' explicit prompts and direct student input, such as students evaluation of the feedback provided, whilst others, such as the mouse position, are recorded without the students' awareness. It is possible that random variables result from the prospective outcome emotions network, which also influence achievement emotions experienced during student interaction with the game challenge. Hence, it is also important to incorporate these variables when applying the algorithms for identifying potential conditional independent or dependent relations (CIDRs), PC and Necessary-Path Condition. Figure 5.24 shows PlayPhysics' retrospective outcome emotions network. It is noted that to assess the experienced *achievement emotion* towards the latest outcome, it is necessary to consider the latest state of the random variables employed to infer *activity emotions*, in addition to intentional variables for capturing student perspectives towards the outcome attained, such as the student eagerness for publicising their results. This random variable may be associated with students' feelings of pride or with their enthusiasm for competing and comparing their performance with their peers.

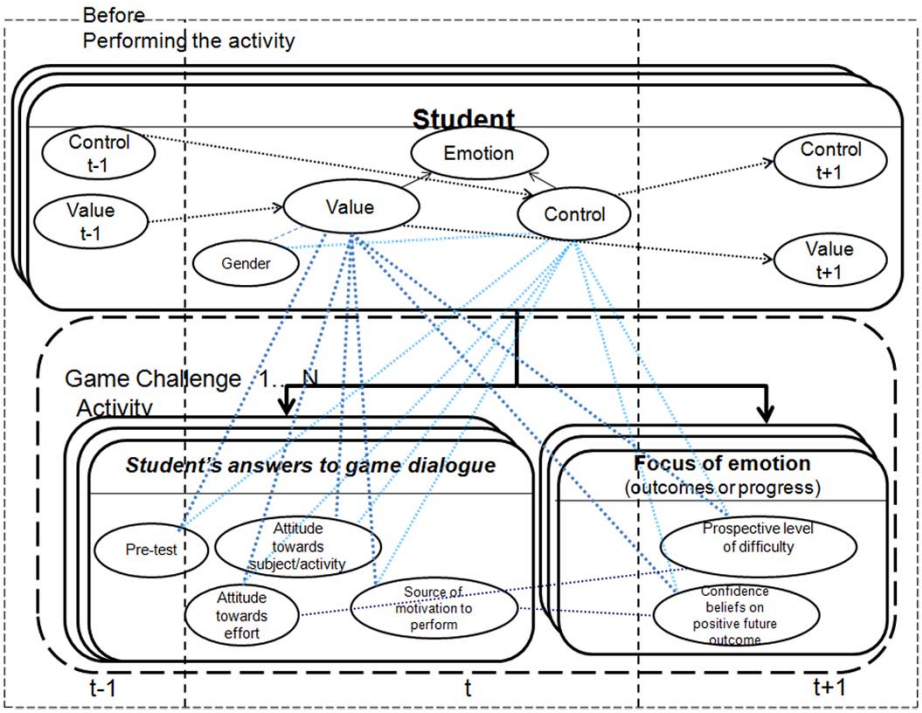


Figure 5.22 PlayPhysics *prospective outcome emotions* network

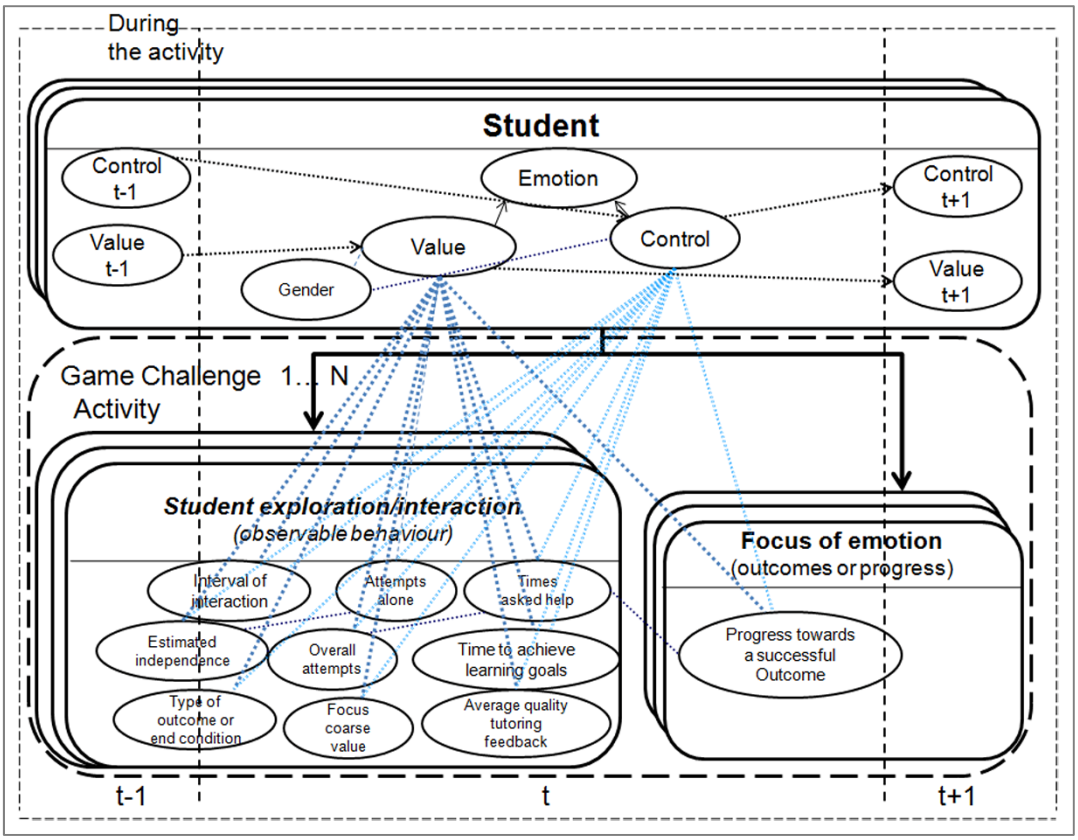


Figure 5.23 PlayPhysics activity outcome emotions network

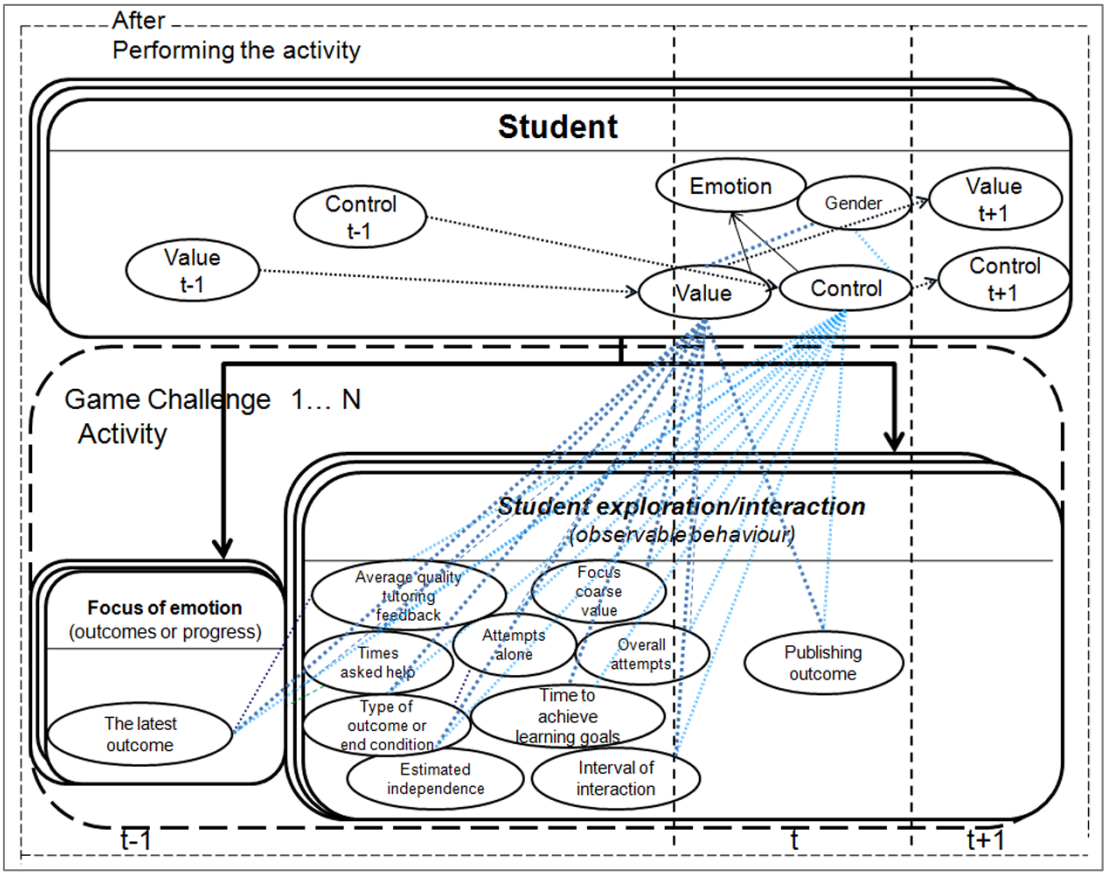


Figure 5.24 PlayPhysics retrospective outcome emotions network

5.5.2 Olympia architecture and software platform

The Olympia architecture integrates GBL environments and ITSs (see Figure 5.25). Olympia provides a semi-open learning environment (Bunt and Conati 2003) where students explore the environment directed by the achievement of specific learning objectives. Olympia is comprised of several modules. The *interface analysis module* filters and selects events that provide information on student interaction and behaviour and which are related to student cognitive and affective states. The chosen events are sent to the behaviour analysis module for evaluation. This process results in evidence forwarded to the student model which includes a *cognitive student model* and *an affective student model*. The *cognitive student model* focuses on identifying student misconceptions and determining student proficiency level, whereas the *affective student model* focuses on identifying student affective or motivational needs and corresponds to our proposed emotional student model discussed in Chapter 4.

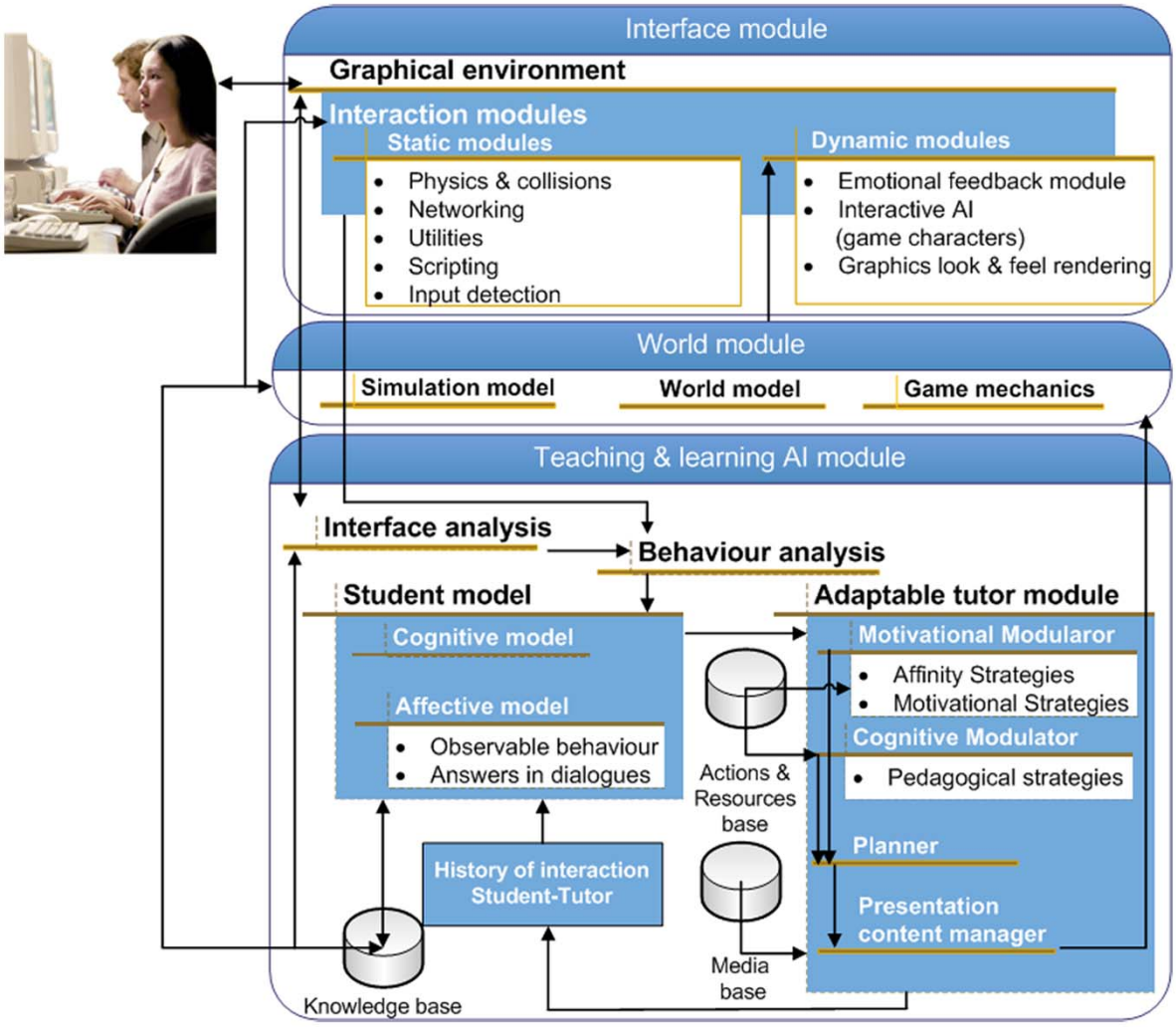


Figure 5.25 Olympia architecture

The diagnosis by the student model is communicated to the *adaptable tutor model*, the main function of which is to select a suitable response according to the information provided

by the *student model module*. The *cognitive modulator* and the *motivational modulator* make suggestions to the planner on the actions and resources to select according to diagnostics received from the *student model*. The planner reviews the suggestions and makes final decisions relating to the display of content. The presentation content manager receives decisions taken by the planner and modifies the game-mechanics accordingly, i.e. action-challenge relation, the *simulation model* and the *world model*. The world model is related to the dynamic interaction modules, which handle content resources that can adapt over time according to identified motivational and cognitive needs. The simulation model is concerned with the representation of physics.

The dynamic interaction modules handle game characters or learning companions (*interactive AI module*), colours or sounds (*emotional feedback module*) and graphics or animations (*graphics and look & feel rendering module*). The *input detection module* is one of the static interaction modules and its main goal is to sense and process input. The *networking module* transmits data across the network, since Olympia is an architecture that supports online gaming and learning. The *Physics and collisions module* includes all the physics and maths controlled objects, which endeavours to enhance the level of realism of the simulation. The PlayPhysics application is developed with the Olympia architecture, an emotional game-based learning environment for teaching physics at undergraduate level.

5.5.3 PlayPhysics components

Here we describe some of the key PlayPhysics components. We use Model-View-Controller (MVC) (Fowler 2008), which is a design pattern where Model corresponds to information about the domain, e.g. an object or data structure. View characterises the display of the model in the GUI, e.g. JSP or HTML page. Controller handles PlayPhysics changes or transitions.

An example of MVC in PlayPhysics is when a student registers, PlayPhysics requests from the student his/her personal details, such as firstname, lastname, group and password, through *registerStudent.jsp*, the view (see Figure 5.26). When the student submits the data, *RegisterStudent.java*, the Servlet and controller, receives it through HTTP POST and validates it. Potential exceptions or errors are handled, such as empty fields or non-matching passwords (trajectories indicated by the orange coloured arrows). To save data in the database the class ConexionDB.java handles connection to the PlayPhysics database (the model) with Java Database Connectivity (JDBC). Finally, it is signalled that the operation was successful (trajectory indicated by the green coloured arrow). Results of the processing are displayed via HTML. Most of the use cases identified during PlayPhysics design, such as ‘Register semester’, ‘Modify and delete groups’ and ‘Modify personal details’ , are implemented with MVB.

The functionality of the Physics Simulation Model contains ShipController.js, a JavaScript class, which includes the spaceship object definition, i.e. parameters and methods (see Figure 5.27). This class has parameters for the acceleration and velocity in three dimensional displacement (x, y, z) and rotational ($roll, pitch, yaw$) axes (see Appendix F). The ShipController JavaScript class is called by the ChallengeControlller class (Figure 5.27(1)) for defining the specific behaviour according to the game challenge goals and constraints. For example, ChallengeController.js randomly assigns the following constraint variables: initial distance to Alpha Centauri, D, and time to exhaust the fuel, T. However, it only initialises them if the student has not interacted with the game challenge previously or has not achieved the educational game goals. Also, it converts units from the International System of Units (SI) from the French 'Système international d'unités', to units in the 3D world and vice versa.

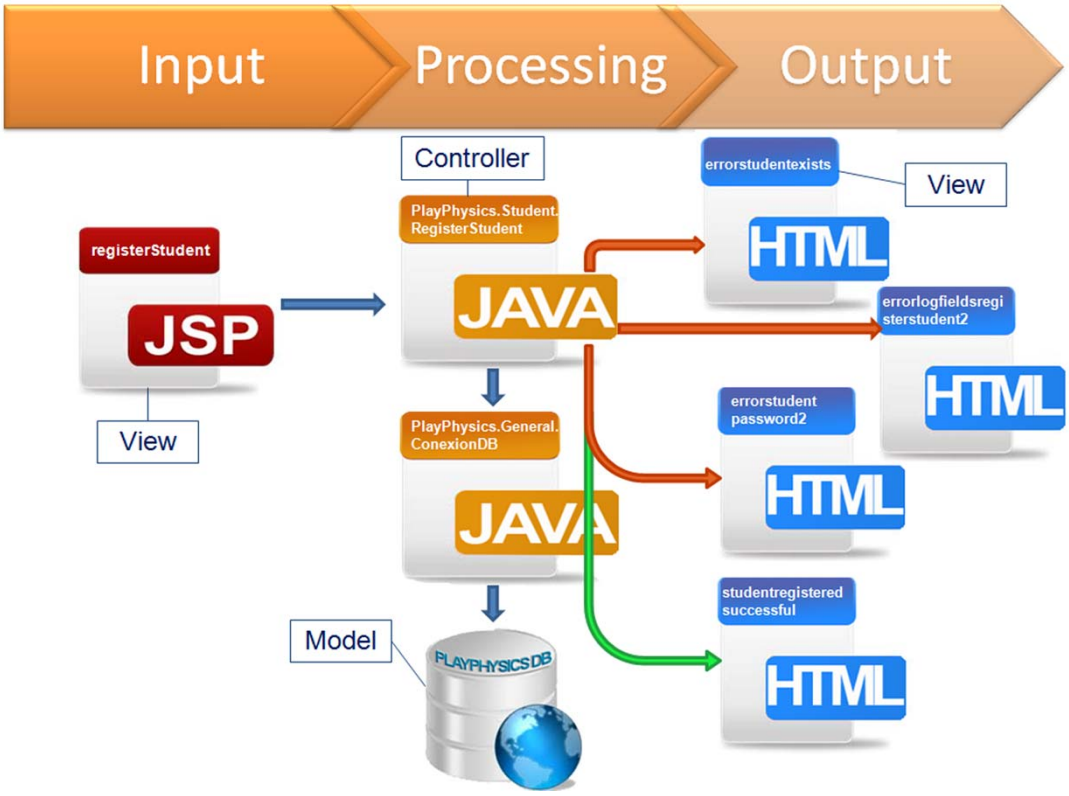


Figure 5.26 Subset of Model-View-Controller (MVC) in PlayPhysics

PlayPhysics' game challenge GUI display is handled by *FirstLevelGUI.js* (Figure 5.27(2)). As a result, *ChallengeController* and *FirstLevelGUI* are two JavaScript scripts that were created to operate together. *FirstLevelGUI* creates an object of the Challenge Controller class for access to its functionality. In a similar manner, *ChallengeController* can request a reference to *FirstLevelGUI* to initialise constraint variables and update the values according to the behaviour expected by the analysis of the physics domain, such as advancing Alpha Centauri at a specific velocity towards Athena or to capture interaction data for the PlayPhysics

database (Figure 5.27(3)). *FirstLevelGUI* can also access the WWW utility to request data recording related to interaction events in the PlayPhysics database (Figure 5.27(4)).

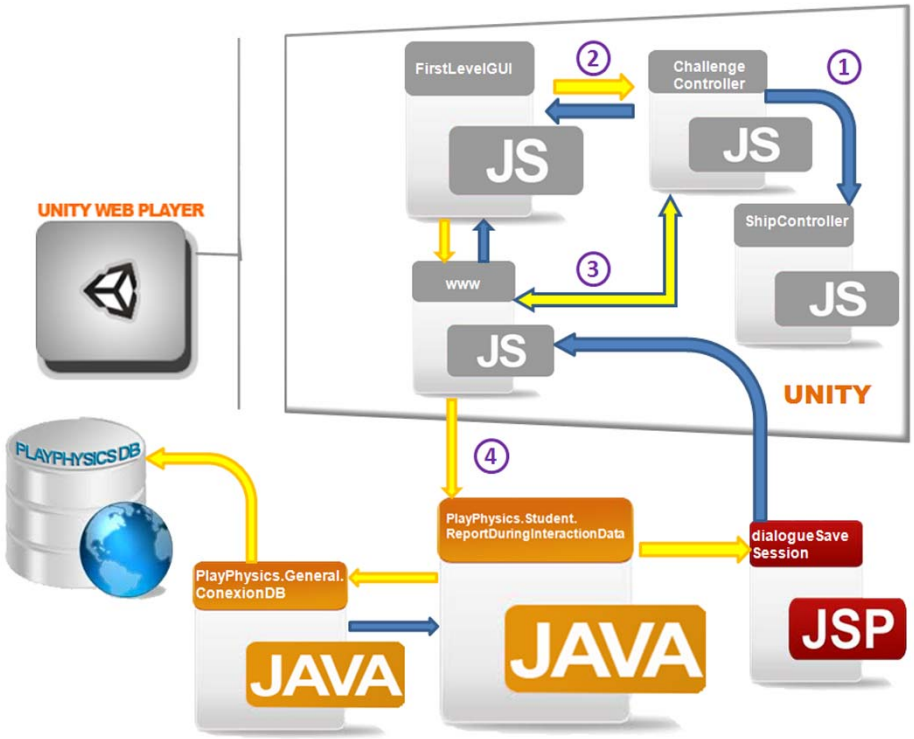


Figure 5.27 JavaScript class interaction for World and Simulation models

5.5.4 Class diagram

An Object-Oriented (OO) paradigm is employed for designing PlayPhysics. A class diagram comprising key classes involved in supporting student interaction with PlayPhysics is shown in Figure 5.28. This class diagram shows class attributes, methods and relationships with other classes. The FirstLevelGUI_PhaseI is a class that controls the graphical display of the game challenge. It updates the variables related to the status of the execution of the game challenge, such as current distance to Athena, the time to exhaust the combustible and the quantity of combustible available. For this reason, FirstLevelGUI_PhaseI contains an object of the ChallengeController class for acquiring this information. FirstLevelGUI_PhaseI also comprises methods, such as setValInitialDistance(), which communicates values, input by students in the game challenge GUI, to the physics simulation model. It also employs data from the current StudentData object, such as gender, to decide which graphical resources must be loaded for giving feedback or displaying player characters. FirstLevelGUI_Phase_I, as with other classes involved in supporting the execution of the game challenge, extends from the GameObject class, which comprises methods such as Awake(), Update() and OnGUI() that are related to initialisation of variables, the implementation of game behaviour and rendering and handling GUI events respectively. .

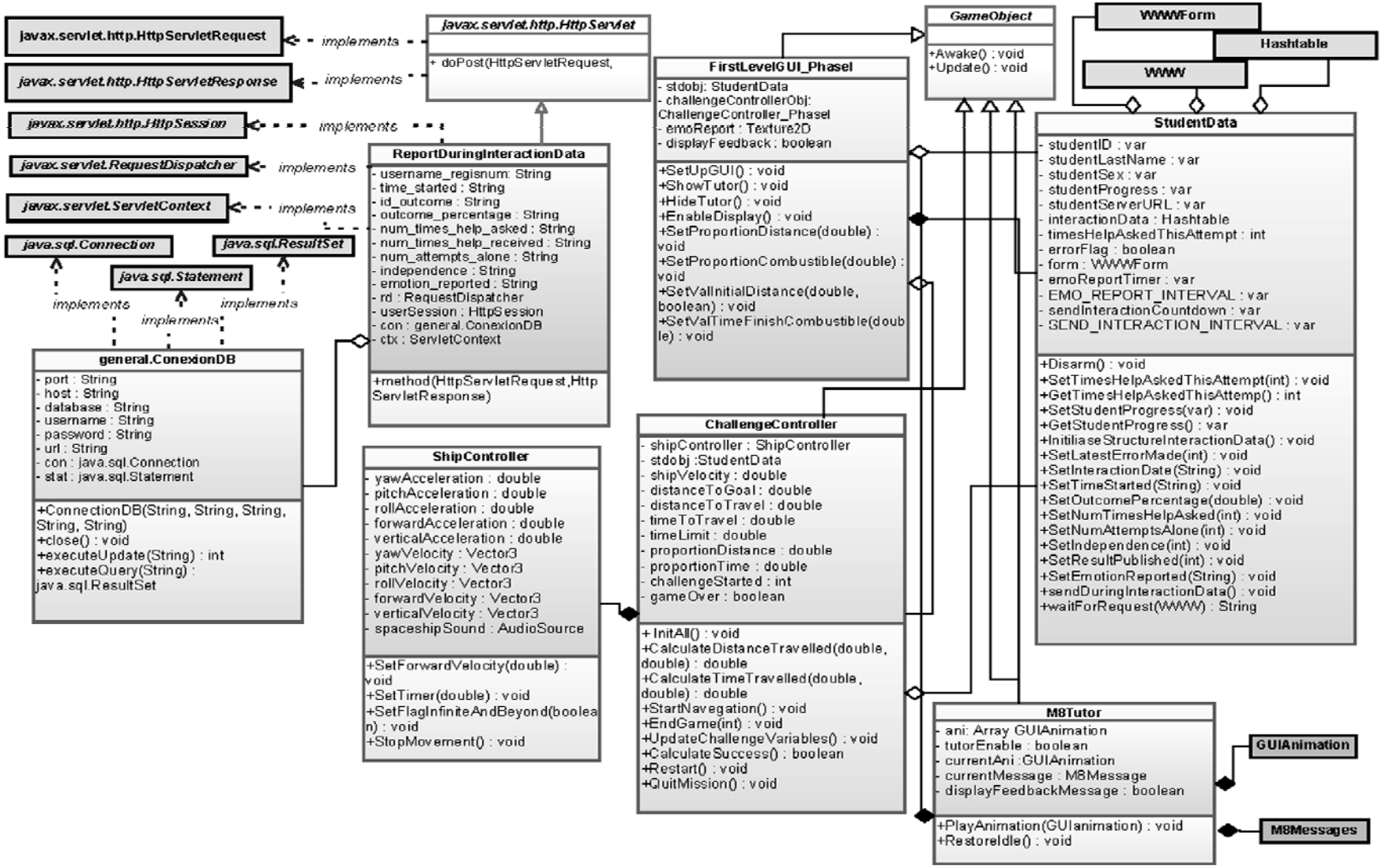


Figure 5.28 Class diagram of key classes for student interaction with PlayPhysics

FirstLevelGUI_PhaseI also controls the display of M8-robot, the PlayPhysics learning companion. Hence, it includes an object of the M8Tutor class. M8Tutor is a class that comprises an array of animations for displaying different pre-defined behaviours, such as blinking or raising hands signalling power and victory. Any pre-defined animation can be played using the PlayAnimation() method and can be replaced by the idle animation with the RestoreIdle() method. Each time that an animation is selected a corresponding written message is also chosen. These messages are handled by the M8Messages class.

The ChallengeController class receives values set by students and commands related to initialising execution of the physics simulation model, stopping and restarting it. This class executes all calculations related to assessing student actions to quantify student knowledge. It uses the CalculateDistanceTravelled and CalculateTimeTravelled methods for this purpose. The ChallengetoController class also handles recording of the interaction variables, for completely defining the proposed emotional student model. Hence, it is capable of accessing the StudentData object that records this data. The SendDuringInteractionData() method sends interaction data through HTTP POST by calling ReportDuringInteractionData, which uses a ConexionDB class object to save data in the PlayPhysics database.

The ChallengeController class includes a ShipController object that receives constraint variables of the challenge and values set by the student for acceleration and initial velocity. Based on the physics behaviour signalled by ChallengeController, ShipController translates these variables into the behaviour that must be observed in the game world. For example, if ChallengeController communicates the value of $1500 m/s$ for the initial velocity to ShipController, this transforms this value into $150 game\_units/s$, so in this case the real world translates into a fraction of $\frac{1}{10}$ in the game world.

5.5.5 Entity-Relationship (ER) diagram

The entity-relationship (ER) diagram or semantic database model of PlayPhysics describes the dependencies between tables and data entities (see Figure 5.29). The tables *students* and *groups* are where most of the dependencies converge and as a result may be considered core PlayPhysics functionality. The Table GSR_Sensor_Measures maintains all the GSR raw measures acquired per student per second with a biofeedback device. All measures are timestamped for different game challenges (Interaction_FirstLevel_Phase I table) or their overture, i.e. cut-scene containing a game-dialogue (FirstDialogue table). The Table QualitativeEval_PlayPhysics contains student responses corresponding to the qualitative assessment of PlayPhysics GBL environment, which may give enhanced insight of student experience of achievement emotions.

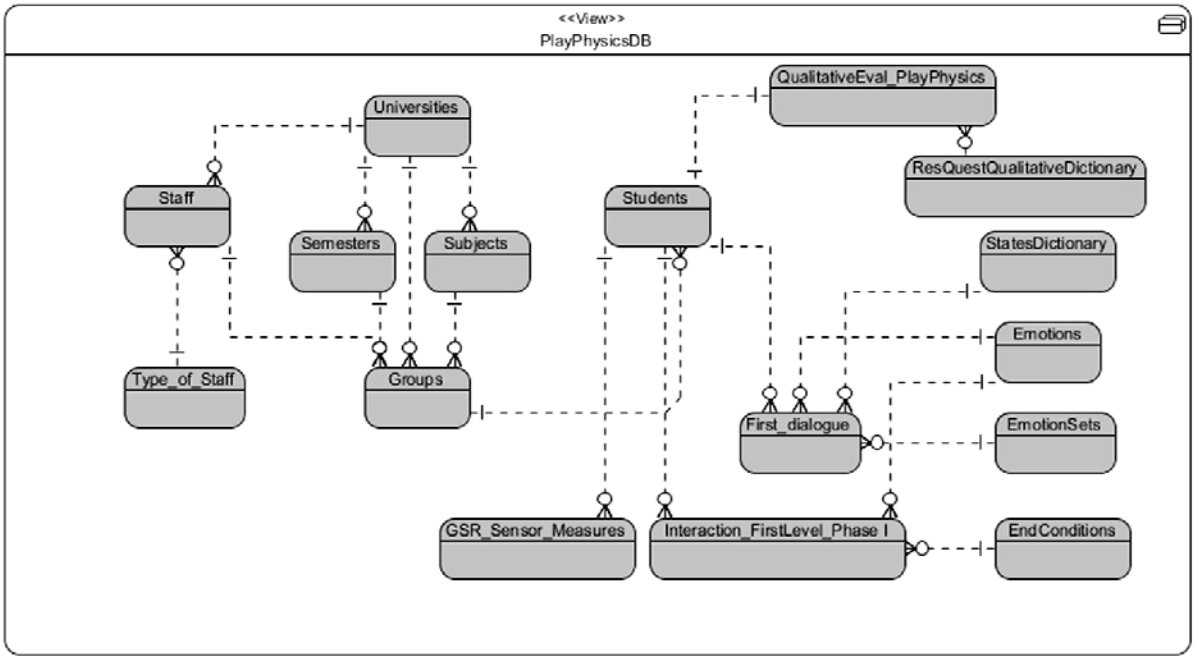


Figure 5.29 PlayPhysics database entity-relationship (ER) diagram

5.5.6 PlayPhysics GUI design

The GUI of the first game challenge in PlayPhysics is created with two views: (1) the cockpit view and (2) the external view. Students can change between views with the F1 and F2 keys. The external view shown in Figure 5.30 enables students to visualise their distance from Athena without being obstructed by the controls. The cockpit view shown in Figure 5.31, is where all the controls related to the operation and navigation of Alpha Centauri, direct-input and self-reporting tools are available. This is the view students are expected to spend most of their time with.

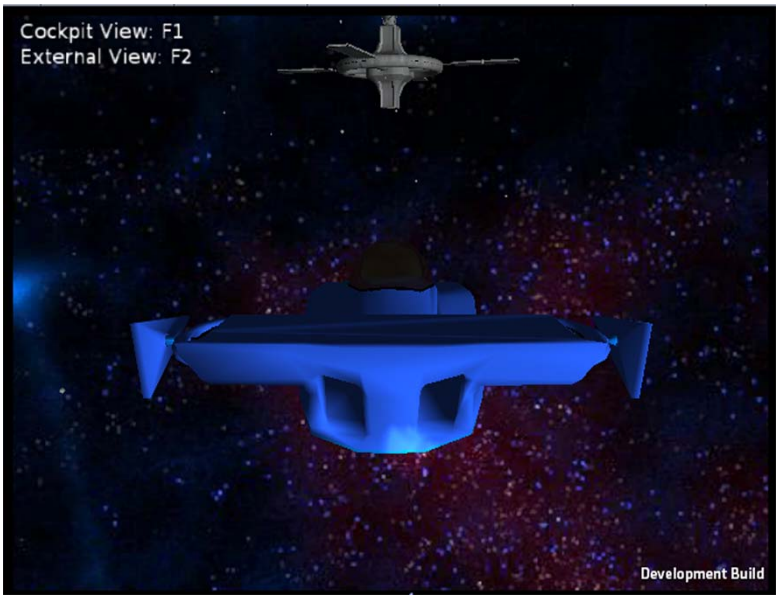


Figure 5.30 External view of PlayPhysics game challenge

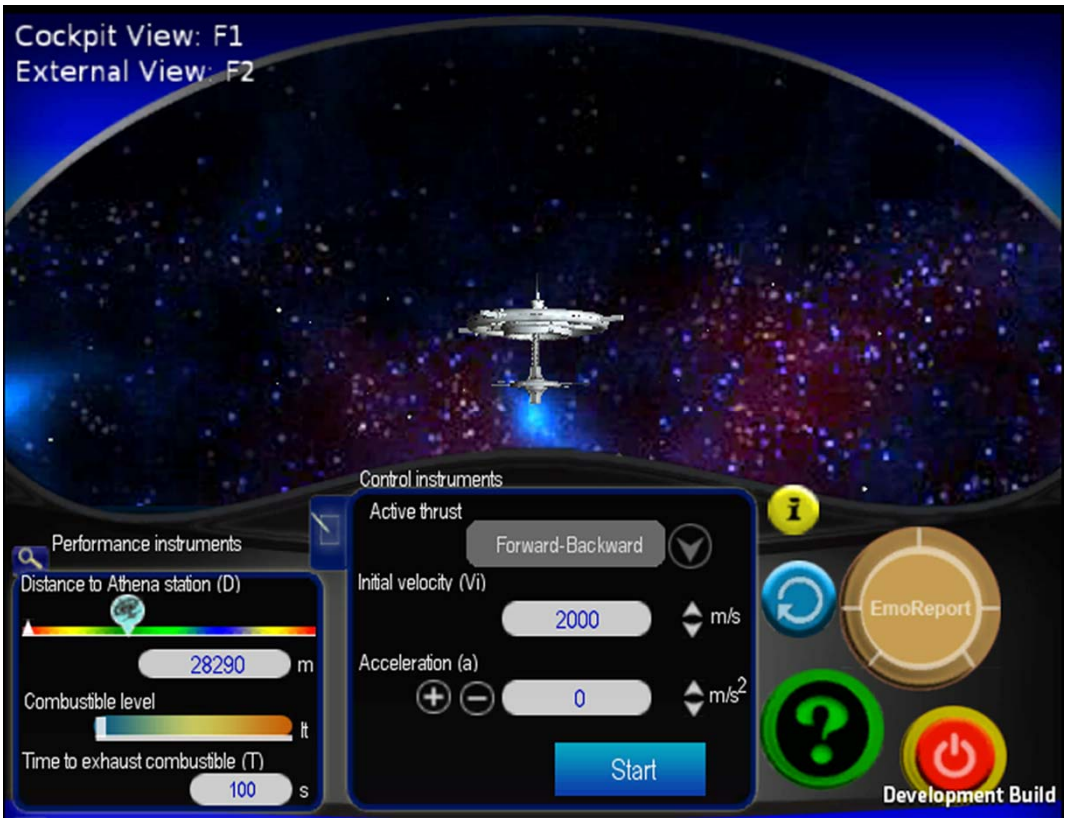


Figure 5.31 Cockpit or internal view of PlayPhysics game challenge

5.6 Summary

This chapter discussed the design of PlayPhysics, an emotional GBL environment for teaching physics at the undergraduate level, in which our computational model of student emotions will be incorporated. For requirements analysis of PlayPhysics a survey was conducted with lecturers and students at undergraduate level giving or undertaking respectively an introductory physics module. The analysis and results of this survey informed the definition of learning goals and design features of PlayPhysics.

The manner in which our emotional student model is populated for PlayPhysics was described. The key goal is to acquire the data to validate this emotional student model, which can reason about emotion using contextual variables. Therefore, it will have the potential to reach a larger population of students via the web. The Olympia architecture, which combines ITSs and game-based learning environments, was also described. Olympia is comprised of several modules, some of which enable the capture of required bio-feedback variables for recognition of emotions..

This Chapter also discussed the functional requirements of PlayPhysics in an intuitive manner from the user perspective, e.g. with written natural language, and on a more formal level, e.g. through use cases, component structures, class diagrams and E-R diagrams. The domain knolwedge of the application domain of PlayPhysics was also de-

tailed. PlayPhysics will quantitatively assess student knowledge through a series of production rules that for identifying student misconceptions. The next chapter discusses the implementation of PlayPhysics.

Chapter 6: Implementation of PlayPhysics

This chapter discusses the implementation of PlayPhysics and describes the manner in which Olympia is applied to the PlayPhysics case study. PlayPhysics is developed using Java, HTML, 3D Studio Max, Poser, Hugin Lite, LightWave, SpaceScape, Javascript, Cascading Style Sheets (CSS), the Unity Game Engine and XML. The implementation of PlayPhysics' input module is key to the application of our proposed emotional student model using Control-value theory, since it is responsible for acquiring data from student interaction. An interpreter of game dialogues was created in order to automatically generate cut-scenes from XML files. The EmoReport wheel was built to allow students to self-report their emotions at anytime. A Bluetooth Galvanic Skin Response (GSR) sensor was built to record student physiological signals using a LEGO NXT intelligent brick and an 8528 converter cable. The GSR sensor is capable of connecting to a Server/PC with compatible Bluetooth capabilities. A learning companion, M8 robot was created by synchronising animations and pre-defined discourse - in the form of written text - for providing instruction, encouraging students to report their emotions, motivating, praising and conveying affinity by mirroring student self-reported emotion. PlayPhysics' communication aspects are described in conjunction with further details about PlayPhysics' implementation. Examples of PlayPhysics' interaction and execution are also presented.

6.1 Implementation environments

PlayPhysics was implemented in Jakarta Tomcat 6.0.32 (The Apache Software Foundation 2011) using Java web technology - e.g. Java Server Pages (JSPs), Java Servlets and Applets – and other web technologies, such as HyperText Markup Language (HTML), Javascript, CSS and the Unity Game Engine. To create the educational game and expedite the implementation of Olympia's modules, such as 'Physics and Collisions', 'Utilities' and 'Scripting', the Unity Game Engine (Unity Technologies 2011) was used, since it facilitates the creation of 3D, 2D and 2½D online games. Its online nature, its Javascript engine and having a general version available for free to developers were key reasons for selecting this tool. PlayPhysics' code is organised in eight packages, where each package comprises a series of classes with specific functionality:

1. The *general* package comprises classes of common purpose, such as classes that handle the connection to the database through JDBC and utilities that facilitate debugging code. This project uses the library LOG4J by Apache Software Foundation (2011) for this specific purpose.
2. The *listener* package includes classes that implement functionality to support the automation of the database backups each time that a University study period is finalised.
3. The *playphysics* package includes classes that extend from the HTTPServlet class, which support different types of HTTP requests. These classes enable staff and students to log-in to PlayPhysics.
4. The playphysics.administrator package comprises the functionality of all use cases in PlayPhysics belonging to the System Administrator role, such as 'modify subject' or 'register semester'. In addition, it comprises all the classes for starting and ending the Bluetooth connection from the PC/Server side with the GSR sensor. Also, it starts the thread that saves data that comes from the GSR sensor in the database.
5. The *playphysics.lecturer* package includes the functionality corresponding to all use cases in PlayPhysics that belong to the lecturer and Head of Department roles, e.g. 'delete group' and 'modify lecturer'.
6. The *playphysics.student* package comprises classes that implement all the functionality in PlayPhysics corresponding to the Student role, such as 'marking pre-test' and the 'advance tutorial'.
7. The *gsr.sensor* package includes the class that starts the connection in the Bluetooth GSR sensor/LEGO Mindstorms NeXT (NXT) generation intelligent brick and waits for the connection on the PC/server. Also it takes the measurement from the skin every second. The NXT brick comprises a 32-bit ARM7 microprocessor, Flash memory, Bluetooth, support for USB 2.0 and four input and three output ports.
8. The *gsrgraph* package comprises all the classes for accessing the student GSR sensor data saved in the PlayPhysics database and visually performing real-time monitoring of the operation of the GSR sensor while it is connected to a student. The main class extends from Applet and it uses the Java chart library 'JFreeChart' (JFree.org 2011), which facilitates displaying high quality charts.

6.2 Communication aspects (Networking module)

The main reasons for the bidirectional communication between Unity and Java are:

- Unity does not have direct access to the PlayPhysics database, implemented in MySQL 5.1 (ORACLE Corporation 2010), for ensuring security. As a result, any

Unity input data, such as personal data (student sex and name) or past progress in the game, is handled via JSP and passed through HTTP POST to the Unity Web Player using JavaScript.

- In the same manner, if the student progress in Unity has to be saved in the Play-Physics' database or an exception error needs to be handled. These are performed through JSPs and the WWW utility scripting module of Unity to retrieve content from Uniform Resource Locators (URLs).
- If Unity requires the loading of specific content in order to display a cut-scene, this information is saved in the form of XML files. These have to be loaded using the WWW utility scripting module and the corresponding XML file. In the same manner if an exception or error occurs, this has to be noticed and handled.
- The emotional model is a dynamic sequence of BBNs, which are implemented using Hugin Lite. Hugin Lite requires a bidirectional communication with the Java classes, but not directly with Unity.

The manner in which the bidirectional communication between Java and Unity was implemented was deduced using the scripting documentation and examples at Unity Technologies (2011) and the samples of connection with PHP at UnityAnswers forum (Qato - Enterprise Q&A 2011). The two methods for communicating input data to Unity are illustrated in the block diagram comprising the communication components shown in Figure 6.1. The first method indicated by the sequence of arrows and numbers in green, starts when the Play-Physics educational game is launched by the student. As a result, *web.jsp* is loaded and it requests all the necessary data that may be required by the PlayPhysics educational game to the PlayPhysics database through the *ConexionDB.java* (Figure 6.1(1)).

Afterwards the data is returned to the *web.jsp* by the *ConexionDB.java* and saved in hidden HTML fields (Figure 6.1 (2)). In parallel, the Unity web player is loaded by the web.jsp (Figure 6.1(3)). When this is fully loaded, it executes StartMenuGUI.js and calls that script's Awake method, which is employed to initialise the 'stdobj' object that contains the data for the student. The 'Awake' method calls the function in the *web.jsp* named 'submitDataUnity' (Figure 6.1(4)), see Appendix E. The 'GetUnity().SendMessage' function obtains a reference to the Unity web player, which receives a message to execute the method 'SetStudentVars' (Figure 6.1(5)), which is located in *StartMenuGUI.js*. The latter is embedded in the 'Main Camera' game object in the scene, and receives the initialisation parameters 'unityParams'. This data is employed to initialise the *StudentData* object (Figure 6.1(6)).

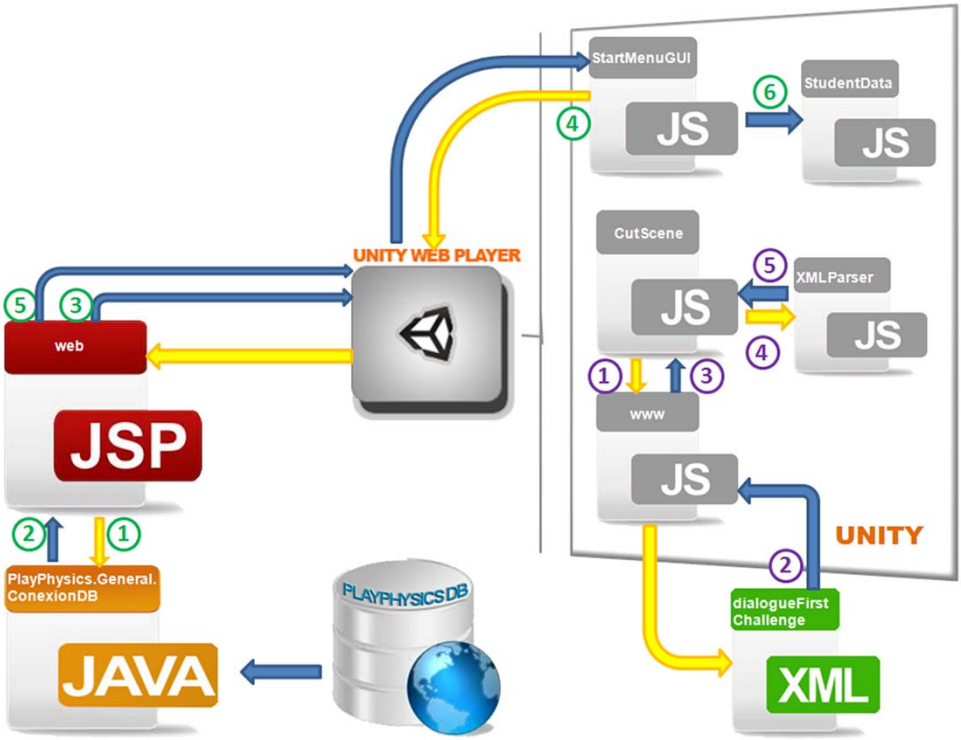


Figure 6.1 Two methods implemented for communicating input data to Unity

The WWW scripting functionality of Unity is illustrated in Figure 6.2 using the arrows and sequence of numbers in purple. For example when the *CutScene* script requires the dialogue corresponding to the first cut scene, this is acquired by executing the WWW function (Figure 6.2(1)) and loading the *dialogueFirstChallenge.xml* (Figure 6.2(2)). WWW handles potential errors and passes the received data to *CutScene.js* (Figure 6.2(3)). If the execution was successful, *CutScene.js* has to parse the XML using *XMLParser.js* (Figure 6.2(4)) in order to make sense of the data provided and display it appropriately in the scene (Figure 6.2(5)), see Appendix E.

There is a need to transfer information from Unity to Java in order to save information in PlayPhysics' database or perform an operation in Java that cannot be performed in Unity, such as propagating evidence in a BBN. This method of communication is also exemplified in Figure 6.2. *CutScene.js* requires saving student data corresponding to his/her progress. To achieve this goal the WWW utility (Figure 6.2(1)) in the 'SaveSession' function, is used to pass data via HTTP post to the Servlet, *RegisterFirstDialogue.java* (Figure 6.2(2)). The latter requests the data and uses the class *ConexionDB* in order to create or update it in the PlayPhysics database (Figure 6.2(3)). *ConexionDB* notifies *RegisterFirstDialogue* about any exception or error that may occur during the execution of this operation, or if it was successfully executed (Figure 6.2(4-5)). Finally to register the outcome of the operation to Unity, *RegisterFirstDialogue* communicates this outcome to *dialogueSaveSession.jsp*, which is sent to the initial WWW request (Figure 6.2(6)). WWW provides the data to *CutScene* as a complete

HTML file (Figure 6.2(7)), which has to be passed to the *HTMLParser* in order to parse it and extract the meaningful information (Figure 6.2(8)). Finally the *HTMLparser* returns the information to *CutScene.js* (Figure 6.2(9)), see Appendix E.

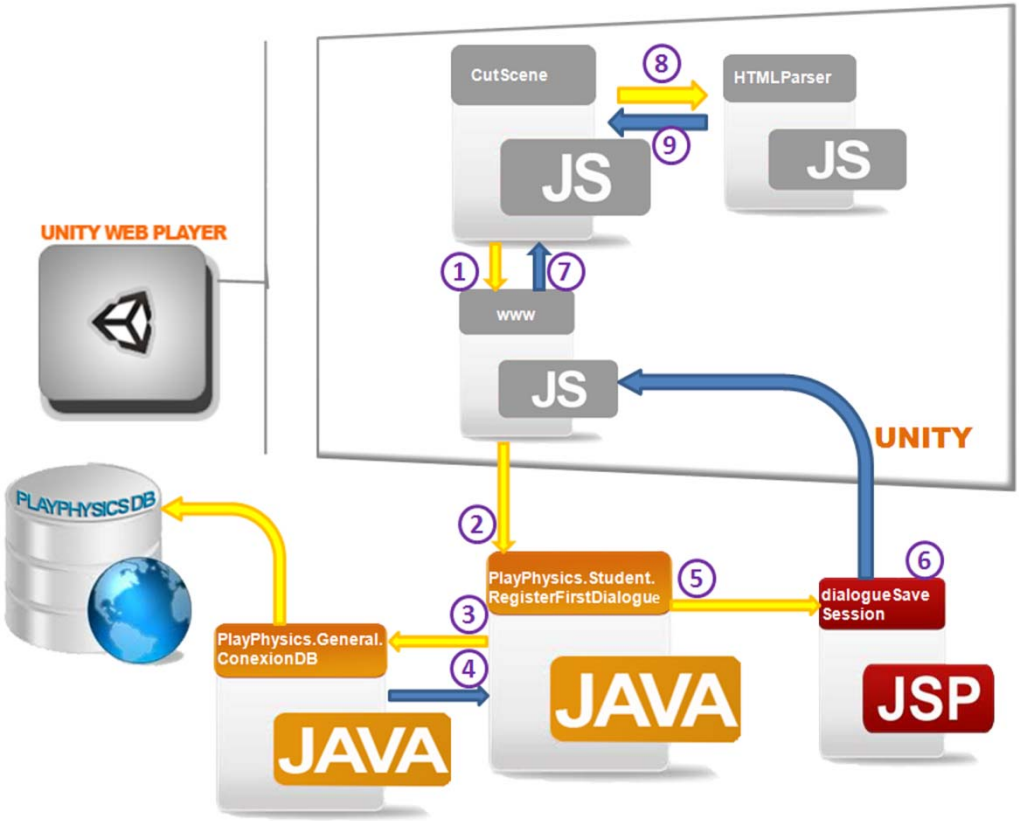


Figure 6.2 Communication from Unity to Java through HTTP POST

In Figure 6.2 a WWWForm is created to send all the parameters through HTTP POST. These parameters are added using the 'AddField' function. This function receives the name of the parameter and its content. The 'yield' instruction allows the WWW function to run as a co-routine and waits for a response before progressing. The 'method' function receives data from Unity through HTTP POST using the 'getParameter' method and saves the data in the PlayPhysics database. It then receives and conveys the result of the database operation by creating a *ConexionDB* object. On obtaining the Servlet context and using the 'getRequestDispatcher' method, the result is redirected to *SaveSession.js*, which includes the outcome of the database operation and is sent in response to the WWW utility execution in Unity. *SaveSession.jsp* transforms the result into an HTML file that is received by *CutScene.js* through WWW. This file has to be manually parsed. If there is no error, Unity advances the student to the next game scene.

6.3 Input detection and interface analysis modules operation

PlayPhysics' design requires that students use the mouse and keyboard to interact, since these are readily available input devices. More advanced peripherals, such as the WII remote control or the PlayStation Move controller, may not be available to all our participants, since only 30% of our students play games 'often' and a further 30% only play games 'occasionally' and we cannot guarantee that even these groups possess the same game consoles. In addition, the use of more advanced peripherals conflicts with our goal of deriving an emotional student model that can identify emotion from contextual and interaction variables related to student behaviour during common online interaction and gaming, with the aim of targeting, and potentially benefiting, a larger population of students. Hence, in order to meet the research goals this section explains how the selected variables, introduced in Chapter 4 and related to student belief and behaviour, are monitored and how these influence the GUI design of the game challenge. As explained in Chapter 4, to reason accurately about emotions an appropriate mixture of variables that are derived from the context as well as variables that depend on student input or self-reporting are required.

6.3.1 Acquiring student belief from PlayPhysics' cut-scenes

One of the ways of acquiring student self-reporting information about their beliefs and attitudes is using game dialogues or cut scenes where questions are presented. From the variables selected for the time frame *before* in Chapter 4, question items were created. However, to display this kind of dialogue and reduce the development effort, a method where the dialogue is passed to Unity in the form of an easily modifiable XML file was designed. It is important to observe that the required graphic design material, such as PNG, JPG or TIFF images, exists under the Unity → Project → Standard Assets folder, so it can be included and displayed. PlayPhysics' player characters were created using LightWave (NewTek 2012). Their textures were created using Adobe Photoshop CS4 (Adobe 2012) and finally applied using 3D Studio Max 9 (Autodesk 2012). These characters were also animated using Poser 8 (Smith Micro Inc. 2010), see Appendix C.

An example of the structure of the XML file received by Unity is shown in Figure 6.3. Each dialogue session is divided by a set of opening and closing *session* tags, e.g. <session></session>. All the content between both tags corresponds to the same dialogue session. The content specified via *preamble* and *appendix* tags are presented on their own screen, while the content defined in *question* and *answer* tags is presented in the same screen. The tags *preamble* and *appendix* are used to display content before or after question-answer blocks if any. The *question* tag is employed to define a question and it is possible to provide up to six potential answers by using the *answer#* tag, where # is a number be-

tween 1 and 6. This file is parsed and a fragment of the code that interprets it in Unity is shown in Figure 6.4.

```
<?xml version="1.0" encoding="utf-8"?>
<dialogue>
        <session>
                <preamble val="This is ground control making contact with the Alpha centauri (spaceship) over...." />
                <appendix val="This is the First Lieutenant from Alpha centauri over ..." />
        </session>
        <session>
                <preamble val="Hi Lieutenant, you are ready to lift off and travel to Athena space station. As you know this
                mission is top secret." />
                <question val="We assigned you to it, since your mentor contacted us and told us that you were the best
                student and you are an excellent Physicist. Is it true?" />
                 <answer1 val="Yes, it is true. I always have found physics interesting." />
                <answer2 val="Well... I am neither excellent nor very bad, but I can give it a try." />
                 <answer3 val="Well... I do not consider myself an excellent Physicist, especially since I think it is a very dif-
                ficult subject, but I can give it a try." />
                 <appendix val="What do I have to do?" />
        </session>
...
         <session>
                <preamble val="Ok, the first challenge involves positioning your spaceship just below Athena space station.
                You will have to dock your spaceship with Athena, which is orbiting around the Sun somewhere between
                Mars and Jupiter." />
                <question val="Which emotion best represents your emotional state before starting the first mission?" />
                <answer1 val="Anticipatory Joy - If you feel that the task will be very easy and you are certain to succeed."
                image="Joy" />
                 <answer2 val="Hope - If you feel you might succeed." image="Hope" />
                <answer3 val="Anxiety - If you feel you might fail." image="Anxiety" />
                <answer4 val="Anticipatory relief - If you thought that the task would be difficult, but at the end you could
                handle the situation successfully." image="Relief" />
                <answer5 val="Hopelessness - If you believe you are certain to fail." image="Hopelessness" />
                <answer6 val="Neutral - If you do not feel any emotion." image="Neutral" />
          </session>
...
</dialogue>
```

Figure 6.3 Code showing the structure of the XML file received by Unity

As can be observed in Figure 6.4, the 2D graphical interface of Unity is displayed using the 'OnGUI' function. GUI.Labels are employed to display images and dialogue on the screen, whilst a GUI.Button is employed to enable the student to advance the dialogue at their own pace. The fragment of code that displays preamble tags drives the dialogue of Captain Damien McCarthy from the NASA space centre. This script is part of the Graphics look and feel rendering module.

```
function OnGUI()
{//Checking if we have a valid GUI Skin Object
	if(gSkin)
		GUI.skin = gSkin;
	else
		Debug.Log("StartMenuGUI: GUI Skin object missing!");
	//Setting the labels
	var nasaStyle : GUIStyle = new GUIStyle();
	nasaStyle.normal.background = nasa;
	var maleAstronautStyle : GUIStyle = new GUIStyle();
	maleAstronautStyle.normal.background = maleAstronaut;
	var femaleAstronautStyle : GUIStyle = new GUIStyle();
	femaleAstronautStyle.normal.background = femaleAstronaut;
	var speakStyle : GUIStyle = new GUIStyle();
	speakStyle.normal.background = speak;
	displayGUI();
	if(showDialogueCaptain=="true")
	{
		//Rect is a 2D Rectangle defined by x, y position and width, height
		GUI.Label (Rect((Screen.width*.10),(Screen.height*.05),64,64),"",nasaStyle);
		GUI.Label (Rect((Screen.width*.10),(Screen.height*.0),256,64),"Commander Damian
		McCarthy");
		GUI.Label  (Rect((Screen.width*.20),(Screen.height*.05),395,captainHeight),"",speakStyle);//Speaking  bub-
		ble corresponding to the Captain
		GUI.Label
		(Rect((Screen.width*.20)+45,(Screen.height*.05)+10,325,captainHeight),dialogueCaptain,"dialogueStyle");//
		Speaking bubble corresponding to the Captain
		//Next Button Captain
		if(showButtonNextCaptain=="true")
		{
			if(GUI.Button(Rect((Screen.width*.20)+375, (Screen.height*.05),35,25), "",
			"buttonStyle"))
			{
				showDialogueCaptain = "false";
				showButtonNextCaptain="false";
				state="2";
			}
		}

	}
…
}
```

Figure 6.4 Fragment of code in Unity that displays preamble tags

Screenshots illustrating PlayPhysics' dialogue content are shown in Figures 6.5 to 6.8. The student advances the dialogue at his/her own pace using the arrows that are displayed only when the content of a preamble or appendix tag is displayed on the screen. The arrows change from blue to green when the student positions the mouse over them. Generally a preamble tag will correspond to a non-player character, while an appendix tag will correspond to the student or player character. Figure 6.8 demonstrates the display of a preamble tag that hints at the context the student should be mindful of while answering the question,

which in this case is an assessment of student confidence belief, see Figure 6.7. The answers also change from blue to green when students hover the mouse over them.

Figure 6.5 Screen shot showing the display of a preamble tag

Figure 6.6 Screen shot showing the display of an appendix tag

Figure 6.7 Preamble preparing the context for student answers

Figure 6.8 Question block to assess student confidence beliefs

6.3.2 Acquiring and monitoring PlayPhysics' interaction variables

This section illustrates the manner in which the capturing of observable variables, comprising PlayPhysics' emotional student model and examined in section 5.5.1, is conducted. The observable variable *interval of interaction* is acquired while students are interacting with the game challenge without his/her awareness. It is obtained by subtracting the time that the student game session started from the time that the student game session ended. Both of these times are recorded in PlayPhysics' database in the format 'hh:mm:ss', corresponding to hours, minutes and seconds. However, the interaction interval is transformed and reported in seconds, see Figure 6.9. All the observable variables are kept in the StudentData object, which is updated whenever an interaction event happens and is saved in PlayPhysics' database every time the student self-reports his/her emotion.

```
var startTimeSecs :int = (parseInt(startTime[0]) * 60 * 60) + ( parseInt(startTime[1]) * 60) + parseInt(startTime[2]);
var endTimeSecs : int = (parseInt(endTime[0]) * 60 * 60) + (parseInt(endTime[1])  * 60) +  parseInt(endTime[2]);
var totalTime : int = endTimeSecs - startTimeSecs;
SetInteractionIntervalSeconds(totalTime);
```

Figure 6.9 Code for calculating the interval of time spent on the game session

The temporal variable *time to achieve the learning goal* records the total time spent by students working with the game challenge before successfully achieving the game result for first time. A similar calculation to the one shown in Figure 6.9 is employed to document this value. However, this value is only saved in PlayPhysics database in the format 'hh:mm:ss'. The random variable *outcome* corresponds to the final result of the game challenge, which can receive values from 0 to 100, as discussed in section 5.3.1. This is the quantitative evaluation of student actions provided by the cognitive student model, the implementation of which is discussed later.

The random variable *times help asked* depends on direct student input. It is the result of asking M8-robot for a hint. This is achieved by clicking on the question mark button in the report panel, see Figure 6.10. M8-robot will provide guidance in the feedback panel. However, M8-robot will only provide a hint when the student has attempted the challenge at least once. Otherwise it responds to the student with the message: 'You have to first attempt the challenge, so that I can identify the conceptual error'. The random variable *average quality of tutoring feedback* also requires student input. When M8-robot provides a suitable hint related to a student misconception, the student may provide an evaluation of its usefulness. Figure 6.11 gives an example of a hint given by M8-robot. In this case, the student assigns to M8-robot's hint a 'Medium' level of usefulness by selecting the corresponding option.

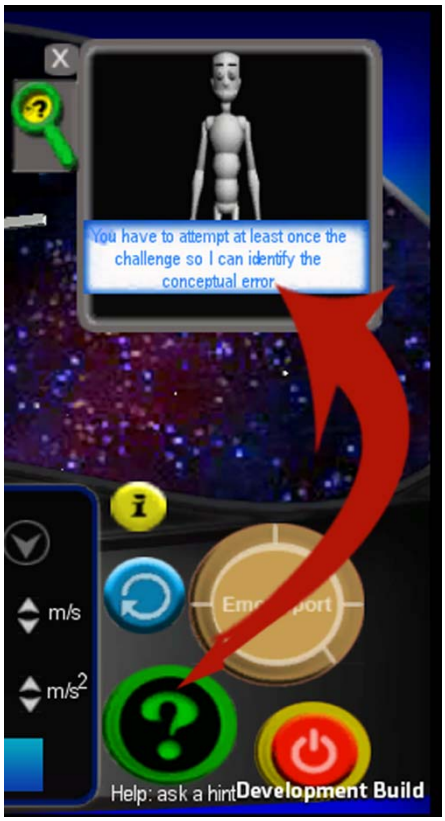


Figure 6.10 Illustration of the student asking for a hint from M8

Figure 6.11 Exemplifying the student assessment of M8 feedback

The random variable *attempts alone* is calculated by adding to this variable each time that the student did not ask for help to *M8 robot* in a trial or exploration. The variable *estimated independence* is obtained by subtracting the *times help asked* variable from the *attempts alone* variable. Hence, if this variable becomes negative, it means that the student is not independent in his/her approach to finding a suitable solution for the game challenge. The random variable *overall attempts* tracks the student effort towards achieving the game challenge goals. It is obtained by adding to this variable for each student endeavour in a student session. Figure 6.12 shows a fragment of the EndGame function employed to report the end condition of the game challenge. Note that the calculations for the interaction variables men-

tioned previously are exemplified in this code section. As can be observed, the results of these calculations are saved in the StudentData object 'stdobj' by executing the Setters, i.e. *mutators*, of these variables. However, the values of these variables can be obtained by invocating the Getters, i.e. *accessors*, of these variables.

```
function EndGame( endType : int )
{
        gameOver = true;

        if (stdobj != null)
        { if (endType>GAME_ENDED_SUCCESS && endType<GAME_ENDED_QUIT)
                {
                        stdobj.SetLatestErrorMade(endType);
                }
                stdobj.SetEndCondition(endType);
                stdobj.SetTimeEnded(System.DateTime.Now.ToString("hh:mm:ss"));
                stdobj.SetTotalAttempts(stdobj.GetTotalAttempts() + 1);
                if (stdobj.GetTimesHelpAskedThisAttempt() == 0)
                {
                        stdobj.SetNumAttemptsAlone( stdobj.GetNumAttemptsAlone() + 1 );
                }
                stdobj.SetIndependence(stdobj.GetNumAttemptsAlone() - stdobj.GetNumTimesHelpAsked());
        ...
        }
}
```

Figure 6.12 EndGame function used to set the finishing cause of the game challenge

The variable *end condition* corresponds to the reason for which the student attempt was finished or signals if the student attempt is still in progress. It can receive values from 0 to 8 and corresponds to a status or kind of end in the execution of the game challenge, see Table 6.1. It is important to know if the student stopped interacting with the game or restarted it, since stopping may be an indication of anger or frustration. This is achieved by clicking on the red button with the 'power off' symbol in the report panel (Figure 6.10). Re-start is an option that allows the student to try again without waiting for the *game over* animation to complete. Re-starting is possible by clicking the blue button with the 'reload' symbol in the report panel, see Figure 6.10).

As discussed in Chapter 3, section 3.7.3, the mouse pointer can be employed to infer user concentration and object of focus (Huang et al. 2012, Wilson 2010). To achieve a measure of student attention, which may be linked to student interest, it was decided to assign a numeric value to certain areas of the game screen. The areas are shown in Figure 6.13 and ordered by the area where it is more likely that the mouse pointer would be located if the student is paying attention (4) to the area which would indicate the student is not paying close attention

(0). If the mouse is not located in any of these areas a value of -1 is assigned to the location of the mouse. The mouse location is recorded with a frequency of 5 milliseconds. The numeric value of the location of the mouse is employed to calculate a coarse value corresponding to the mouse focus (*Focus coarse value* variable), which gives an approximate measure of how interested the student is in solving the case study.

End condition	Description
0	In Progress
1	Success accomplished
2	Misconception: Positive acceleration
3	Misconception: Black out, acceleration larger than $4g$
4	Misconception: Distance travelled is too short from Athena
5	Misconception: Distance travelled is too far from Athena
6	Misconception: Time out, the combustible was exhausted
7	Quit the game challenge
8	Restarted the game challenge

Table 6.1 End conditions of PlayPhysics game challenge

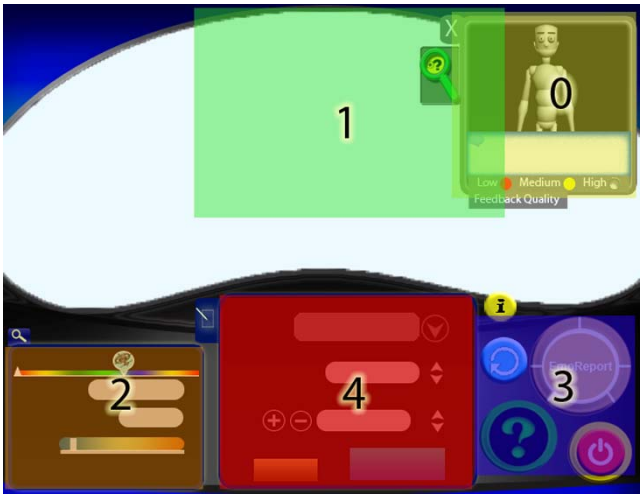


Figure 6.13 Numerically weighted screen areas for categorising the mouse focus

The fragment of code that performs the monitoring of the mouse position and registers its value is shown in Figure 6.14. An array of rectangles, FOCUS_RECTS, was created in order to easily manage these areas. MOUSE_CHECK_INTERVAL is a constant that represents the five millisecond interval corresponding to the mouse focus monitoring frequency. The 'Time.deltaTime' function makes the game *frame-rate independent*. It corresponds to the time in seconds between the last frame and the current frame, i.e. fixed frame-rate. If the mouse is not inside one of the rectangles a value of -1 is recorded, otherwise the index of the rectangular area is recorded. For example, if the mouse is over the feedback panel, where M8-robot is situated, the value assigned is 0 and if the mouse is located over the control

panel, the value assigned is 4. This information is recorded in the Illustration of the numerically weighted screen areas for categorising the mouse focus *Student Object*, stdobj.

```
static public var FOCUS_RECTS : Array = new Array(
                                Rect(458,8,188,190) ,                    // feedback panel
                                Rect(195,4,317,214),                     // view rectangle
                                Rect(2,350,202,145),                     // performance panel
                                Rect(482,290,163,163),                   // report panel
                                Rect(221,297,258,198)                    // interaction/control panel
                               );

private var mouseCheckCounter : float = 0.0f;
static private var MOUSE_CHECK_INTERVAL : float = 5.0f;
...
function Update()
{
        if (tickersPressed == 0)
        {
                tickerVal = 1;
        }

        mouseCheckCounter -= Time.deltaTime;
        if (mouseCheckCounter <= 0)
        {
                mouseCheckCounter = MOUSE_CHECK_INTERVAL;
                var mousePos = Input.mousePosition;
                mousePos.y = Screen.height - mousePos.y;
                // check mouse position in rectangles and send to StudentData
                if (stdobj)
                {
                        var rectIndex : int = -1;
                        for (var i : int = 0; i < FOCUS_RECTS.length; ++i )
                        {
                                if (FOCUS_RECTS[i].Contains(mousePos))
                                {
                                        rectIndex = i;
                                        break;
                                }
                        }

                        stdobj.SetMouseFocused( ( stdobj.GetMouseFocused() == "" ? "" : (stdobj.GetMouseFocused() +
                        ",") ) + rectIndex);
                }
        }

…
}
…
```

Figure 6.14 Fragment of code for frequently monitoring the mouse focus

It is also necessary to enable students to self-report emotional state, either when they feel like reporting it during game play or by request, i.e. when M8-robot asks them to do so. The

latter is necessary, so as to ensure collection of the required data for deriving and evaluating the proposed emotional student model. This is achieved via the *EmoReport* wheel in the report panel (Figure 6.10), which also appears in the game score screen. The *EmoReport* wheel that is included while interacting with the game challenge comprises the neutral emotion or no-emotion and the four emotions corresponding to the *activity achievement emotions* by Pekrun et al. (2007), e.g. boredom, frustration, anger and enjoyment (see Figure 6.15).

The *EmoReport* wheel that is presented along with the game outcome comprises the neutral emotion and the *retrospective-outcome achievement emotions*, since according to Pekrun et al. (2007), these emotions are more likely to occur after achieving an outcome, e.g. *anger, sadness, shame, pride, joy* and *gratitude*, see Figure 6.16. Additionally, PlayPhysics asks students about their affective state at the end of the game dialogue and before the interaction with the game challenge. The emotions that are more likely to be reported at this stage are *neutral* and *prospective-outcome* achievement emotions, e.g. *anticipatory joy, hope, anxiety, anticipatory relief* and *hopelessness* (see Figure 6.17).

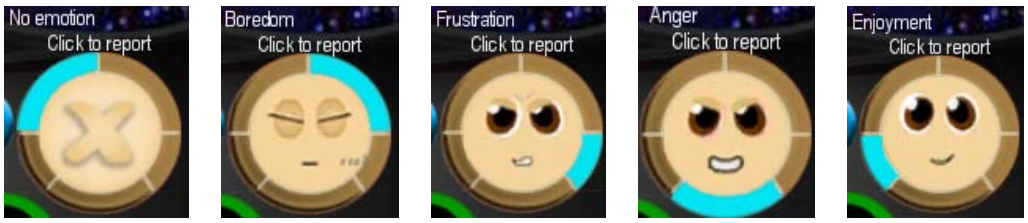


Figure 6.15 Neutral and activity achievement emotions in the EmoReport

Figure 6.16 Neutral and retrospective-outcome achievement emotions in EmoReport

Another variable that depends on direct student input is *publishing outcome*, which is linked to student willingness to allow other students to view the achieved game outcome. At the end of each challenge a list with the ten best scores is displayed (Figure 6.18). It is expected that students may feel the inclination to compete or may feel pride in their achievement. As a result, they may share their outcome. The option for publishing the result only ap-

pears when students have successfully achieved the game goal, i.e. receiving at least a score of 70%. Once students have decided to publish their result by clicking on the corresponding button, PlayPhysics will display the message 'Your result was published' (Figure 6.19).

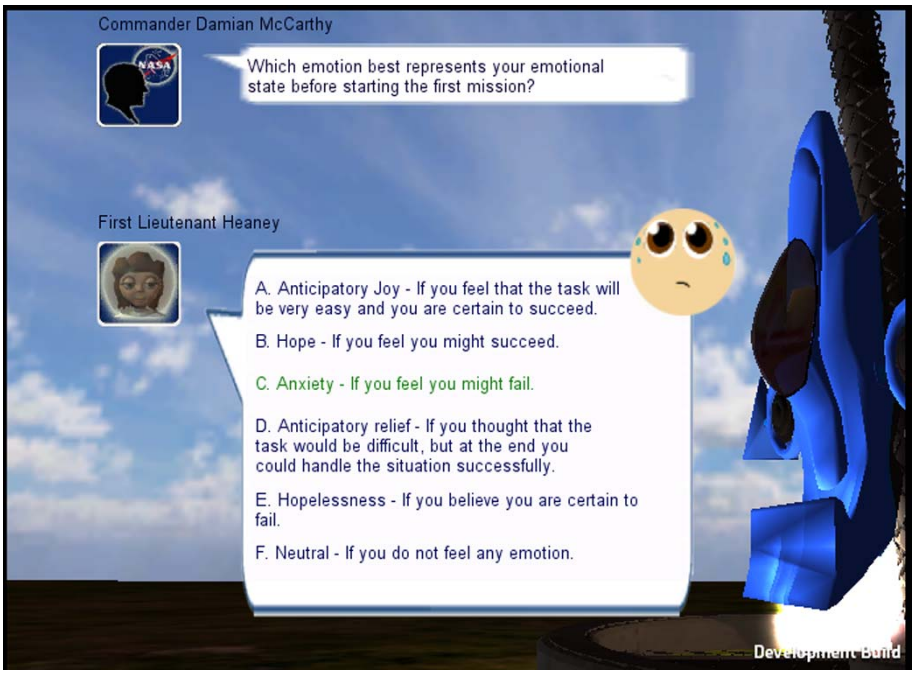


Figure 6.17 Student emotion self-report enquired during the game dialogue

Figure 6.18 Top ten scores selected from the set of students who shared their score

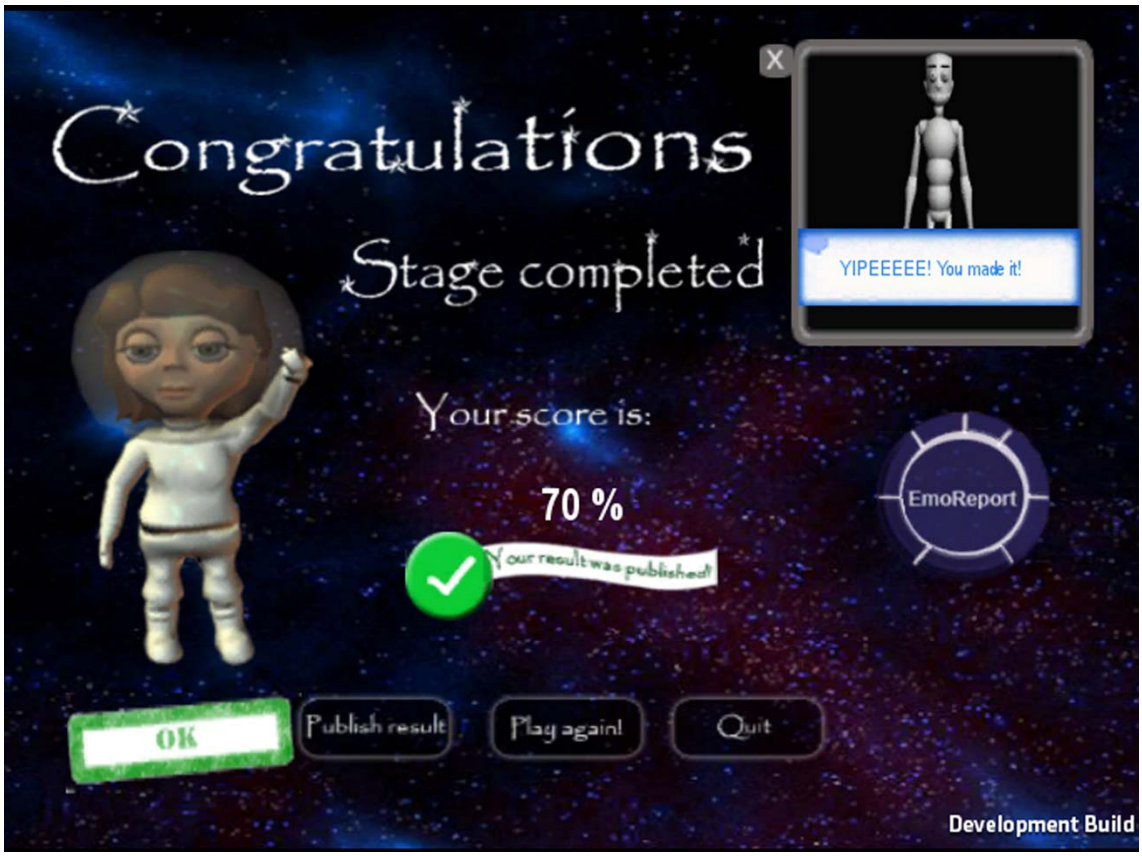


Figure 6.19 Screen shot displaying the student outcome and enabling its publication

6.3.3 PlayPhysics' acquisition and recording of student GSR signal

Another variable that is related to the student affective state and that can be recorded and monitored in PlayPhysics as an on-site GBL environment is the Galvanic Skin Response (GSR). However, the GSR sensor requires a Bluetooth connection (supporting up to 10 m and data rates of up to 3 Mbps). The PC/server Bluetooth adapter employed during testing is a Belkin Mini-Bluetooth adapter, USB 2.0. The sensor was built to the design specification of Gasperi (2010) using the LEGO NXT intelligent brick, foil paper, loop and hoop and an 8528 converter cable.

Gasperi (2010) uses a LEGO Mindstorm NXT program, comprised of blocks, to acquire the raw value of the GSR sensor and display it as a numerical value (between 0 and 1023) in the NXT brick screen. Gasperi (2010) does not communicate the GSR raw value from the NXT brick to the PC. To be effective the sensor needs to communicate in real time with PlayPhysics and provide some way to monitor the connection visually. Therefore, Bluetooth communication, JFreeChart and the LEGO Java Operating System (LeJOS) NXJ (SourceForge 2009) were all employed to achieve this goal. These technologies were selected for compatibility reasons, since PlayPhysics is largely programmed using Java.

In order to work with LeJOS, the firmware of the NXT brick has to be replaced with a Java Virtual Machine. LEJOS has an Application Programming Interface (API) that assists the creation of classes that are converted to binary files and are uploaded to the NXT brick to be executed. LEJOS also comprises a PC API that allows the running of code on the PC in collaboration with the NXT brick using the LEGO Communication Protocol (LCP) in USB or Bluetooth Java streams. The communication between PlayPhysics and the NXT brick is shown in Figure 6.20. In order to enable Bluetooth on the PC/Server, the location of bluecove-*2.1.1-SNAPSHOT.jar* must be added to the System environment variable CLASSPATH. BlueCove is an API for establishing a Bluetooth connection with a JSR-82 implementation that in this case interfaces with the Microsoft Bluetooth stack found on Windows (SourceForge 2008).

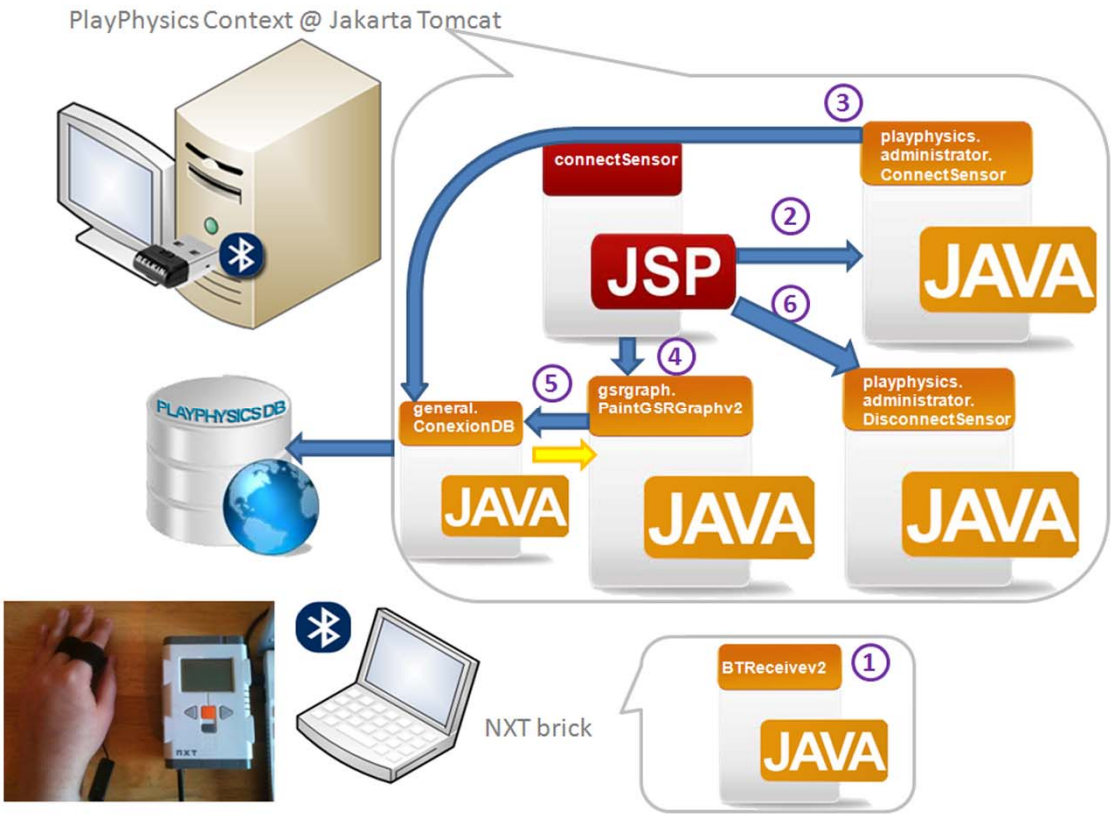


Figure 6.20 Communication between PlayPhysics and the Bluetooth GSR sensor

To communicate, students must already be properly connected to the GSR sensor, which involves connecting the sensor to the index and medium fingers of the least favoured hand, since the favoured hand is used to interact with the PlayPhysics educational game. The GSR sensor is connected to port 1 of the NXT brick. The NXT brick must already be running the program that listens for the connection from the PC side, *BTReceivev2* (Figure 6.20(1)). Then to connect the sensor to a student, who must already be registered in PlayPhysics' databse, the administrator selects the student from a specific group and introduces the name of

the NXT brick to which the connection will be established. The administrator then begins the connection by clicking the 'Connect' button (Figure 6.20(2)). At this point a thread is started that receives the value of the GSR response every second and saves it in the PlayPhysics database (Figure 6.20(3)). While the connection is active, a realtime graph, built using JFreeChart, is displayed in the form of an Applet showing the value of the GSR sensor over time in order to enable monitoring by the administrator (Figure 6.20(4) and Figure 6.21). This displayed value is requested from the PlayPhysics database (Figure 6.20(5)). As a result, there is a delay of a few seconds between the display and the acquisition of the GSR raw value. To disconnect the sensor, the administrator clicks the 'Disconnect' button. At this point, the thread that it is acquiring and recording the GSR raw value stops and the Bluetooth connection is terminated (Figure 6.20(6)).

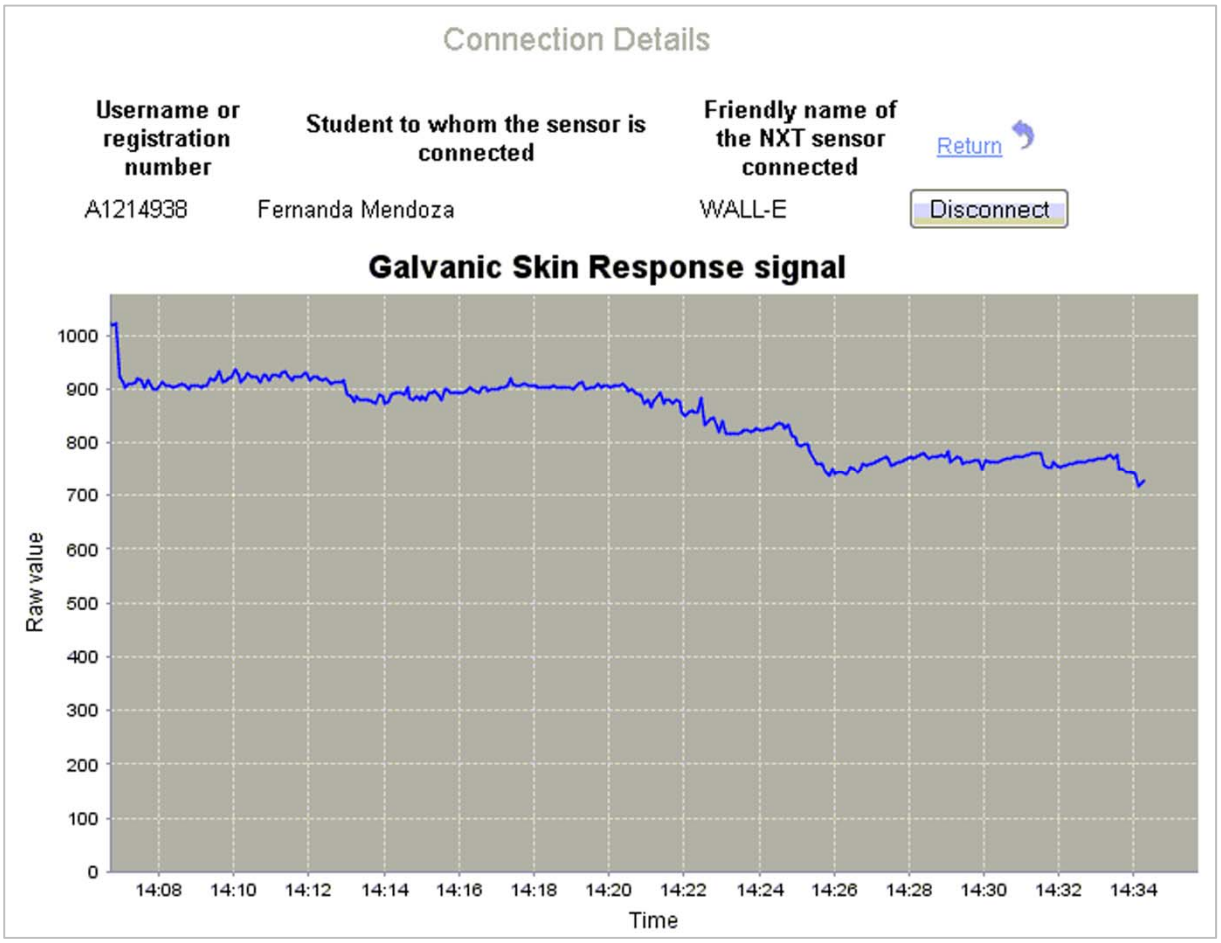


Figure 6.21 Graph of the raw value of the Bluetooth GSR sensor

A fragment of the code corresponding to *BTReceivev2,* which awaits the initialisation of the Bluetooth connection from the Server/PC side, is shown in Figure 6.22. An object of the class *TouchSensor* is created to indicate that the sensor is currently connected to port 1. The variable *sensormeasure* holds the sensor value for sending to the Server/PC side. It is initial-

ised with the value *1023*, since this is the base value provided by the sensor when not connected. The method *readRawValue* is employed to read the value from port 1. *Thread.sleep* is employed to record the measure and send it every second. When *runmeasure* is false the thread and the connection are terminated. *LCD.drawString* is employed to display the sensor value or information about the state of the connection in the NXT brick display screen.

```
public class BTReceivev2 {
        static TouchSensor touch = new TouchSensor(SensorPort.S1);//GSR sensor
        static int sensormeasure = 1023;
        static boolean runmeasure = true;
        static DataOutputStream dos;
         static DataInputStream dis;
        static int value=0;

        public BTReceivev2()
        {
                new a();
        }

        private class a extends Thread{
         public a()
        {
                 start();
         }
         public void run()
         {
                  //Waits for a Bluetooth connection and sends back the data from the GSR sensor
                String connected = "Connected";
                String waiting = "Waiting...";
                String closing = "Closing...";
                LCD.drawString(waiting,0,0);
                NXTConnection connection = Bluetooth.waitForConnection();
                LCD.clear();
                LCD.drawString(connected,0,0);
                dos = connection.openDataOutputStream();//Data send to the computer
                 dis = connection.openDataInputStream();
                 while (BTReceivev2.runmeasure)
                {//Taking the sensor measure
                        BTReceivev2.sensormeasure = SensorPort.S1.readRawValue()+BTReceivev2.value;
                         try {
                                BTReceivev2.dos.writeInt(BTReceivev2.sensormeasure);
                                BTReceivev2.dos.flush();
                                LCD.clear();
                                LCD.drawInt(BTReceivev2.sensormeasure,0,1);
                                //1 second has 1000 miliseconds
                                Thread.sleep(1000);
                            }
```

```
                catch (InterruptedException ie)
                        ...
            }//end while
            try{
                BTReceivev2.dos.close();
                connection.close();
                BTReceivev2.runmeasure = false;
                LCD.clear();
                LCD.drawString(closing,0,1);
              }
            catch(IOException ie)
            ...
        }
...
}
```

Figure 6.22 Code waiting for a Bluetooth connection and transferring the GSR value

6.4 Simulation and world models

The simulation and world models are responsible for ensuring that students experience the effects of the physics phenomena as closely to the real world experience as possible. This involves not only appearance, i.e. look and feel, but also the accurate representation of the phenomena observed. The 3D models of the Alpha Centauri Spaceship and Athena space station were obtained from TurboSquid (2010) for expediting the development process. However, their textures were modified in Photoshop and the 3D models were modified and optimised for real-time rendering using 3D Studio Max. To reduce the quantity of polygons of the 3D models without changing their appearance - reducing the loading time and avoiding slowdown in the game interaction, Polygon Cruncher (Mootools Software 2010) was employed. After installation, Polygon Cruncher is available in the 3D Studio Max interface ready to use. The space skybox employed in PlayPhysics' educational game was created using Spacescape, a freeware tool powered by Ogre3D and Qt. (Geeknet Inc. 2010). The background space and spaceship sounds and music were acquired from Sound-effects-library (2010). The remaining 2D images employed in the Cockpit view were created in Photoshop. The implementation of the Simulation Model used the analysis of the physics domain discussed in section 5.3 and the block diagram of PlayPhysics' components discussed in section 5.5.2 as a reference.

The Unity3D Javascript class including the attributes and methods of the spaceship Alpha Centauri is *ShipController.js.* ChallengeController.js handles and defines Alpha Centauri's behaviour. For example, the initAll function from the *ChallengeController* calls the functions: *SetValInitialDistance* and *SetValTimeFinishCombustible* to initialise the constraint variables of the game challenge with arbitrary values. However, these values are only initialised if the

student has not yet interacted with the game challenge or has not achieved the educational game goals (see Appendix E). ChallengeController communicates the values set by the student in FirstDialogueGUI for the initial velocity and acceleration to ShipController.js and activates Alpha Centauri's movement when students indicate that they want to start the journey towards Athena by clicking on the Start button. ChallengeController also can make the conversion between units from SI to units in the 3D world and vice versa using the functions: *UnitConversion.To Units* and *UnitConversion.To Meters*.

6.5 Behaviour analysis module and cognitive student model

The implementation of the student model and the behaviour analysis module are closely related, since how events or student actions are evaluated assists the diagnosis of student misconceptions and the guidance or feedback that should be provided. In this case study, the cognitive model of the PlayPhysics educational game was not the main focus of this research. Initially we attempted an implementation using DBNs, but it was observed (section 5.3.1) that in this case an implementation using production rules, i.e. if-then-else rules, was more suitable.

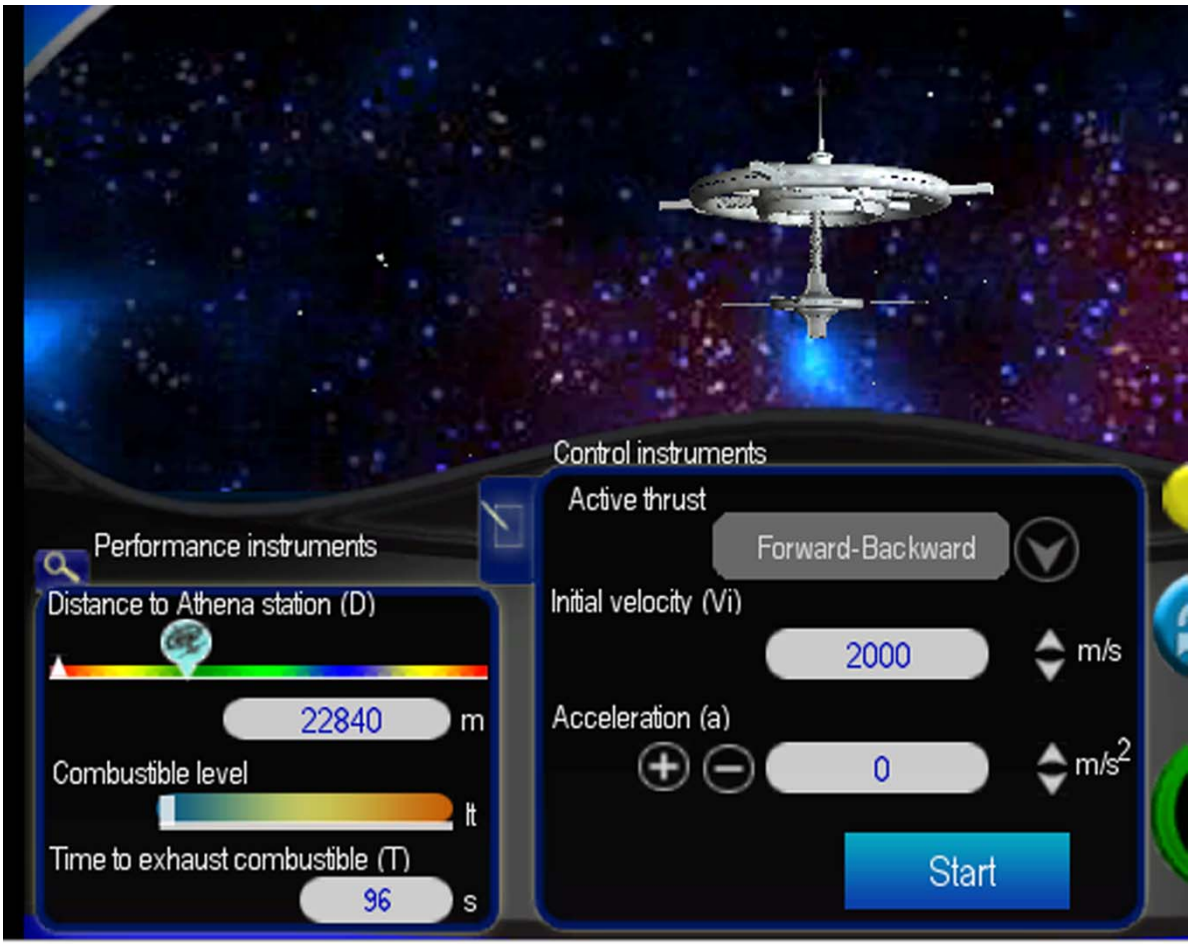


Figure 6.23 Cockpit view of Alpha Centauri: Performance and Control panels

PlayPhysics' game challenges are designed so students can interact from a first person perspective. A game tutorial was created to familiarise students with the functionality of the PlayPhysics GUI (see Appendix F). In the PlayPhysics game challenge, there are two panels on the GUI that must be analysed by students in order to achieve the goal: the 'Performance instruments panel' and the 'Control instruments panel' (Figure 6.23). In addition to the cockpit and external view, the 'Performance instruments panel' assists students in determining how well they are doing. For example, while Alpha Centauri moves towards Athena station, students can see a visualisation of the values corresponding to the Distance to the Athena station, the fuel level and time to exhaust the fuel. This will depend on the values for Alpha Centauri's initial velocity and acceleration. Both variables can be set using the Control instruments panel and then Alpha Centauri starts the trip when the Start button is pressed. The behaviour analysis and student cognitive model assist in determining the quality of the decisions made by students.

The pseudo code describing the behaviour analysis, diagnosis of student needs and production rules implemented as part of the student cognitive model is shown in Figure 6.24. From the physics domain analysis and with the assistance of an Astrophysics expert from ITESM-CCM, nine rules were derived that must be adhered to so as to ensure a successful journey to the space station. In the pseudo code, the variable 'mark' corresponds to the game score achieved, which is given as a percentage. Scores below 70% correspond to non achieved learning/gaming goals or a negative outcome. The direction of the acceleration has a high penalty associated with it. Once student errors are identified, a suitable response is provided by PlayPhysics using implicit clues (suiting more independent learners), or more explicit clues given by the learning companion M8-robot (suiting learners that require assistance to make progress).

```
If (acceleration direction == incorrect or positive)
        mark = 0;
else {// acceleration direction is correct or negative
        if  (a>4g)
        {
           mark = 50;
        }
        else
        {

                    1.   Calculate for this student:
                                 d_S = v_i^2 / 2a                    t_S = v_i / a
                    2.   Take into account the restriction variables T and D

                if (t_s≥T) // The combustible finished
                {
```

```
        mark = 50;
    }
    else {
            //Compare distance
        3.  Calculate relative error of the distance e_d
                e_d = ((d_s - D)/D) × 100
            If (-2≤Rounded(e_d)≤2)
            {// Compare time
        4.  Calculate t_max, t_min and Δt
                a_max = v_imax^2 / 2D  where v_imax = 2000 m/s
                a' = min[a_max, 40 m/s^2]
                t_min = sqrt(2D / a')
                t_max = min[1000/a, T]
                Δt = t_max - t_min
            If (t_min ≤ t_s ≤ t_min + Δt/3)
                mark = 100
            if (t_min + Δt/3 < t_s ≤ t_min + 2Δt/3)
                mark = 95
            else
                mark = 90
            }
            else If (-2> Rounded(e_d)≥-5) or (2< Rounded(e_d)≤5)
                mark = 80
            else if (-5> Rounded(e_d)≥-10) or (5< Rounded(e_d)≤10)
                mark = 70
            else
                mark = 60
        }//end else
}//end else
```

Figure 6.24 Pseudo code of PlayPhysics cognitive student model

6.6 Tutor model and output: Feedback and response of M8 robot

The tutor model uses identified student misconceptions in order to choose the most suitable instructional response. PlayPhysics has two possible types of responses to student misconceptions: (1) implicit or vague feedback and (2) more specific clues. Implicit or vague feedback involves the displaying of images illustrating the issue. For example, Figure 6.25 shows the implicit feedback displayed when it is detected that students have a misconception related to the setting of values for Alpha Centauri's initial velocity (v_i) and acceleration (a) resulting in Alpha Centauri coming to rest too far from Athena station. Table 6.2 shows the im-

plicit feedback created for the PlayPhysics game challenge using the female player character.

Figure 6.25 Implicit response to student misconceptions

The images in Table 6.2 were created using Poser 9 and Photoshop. The provision of more specific clues is given on demand by M8-robot and has the purpose of enabling more independent students to work out the solution by themselves, while supporting students that need assistance to solve a problem. This kind of feedback aims to give a more elaborate explanation of the cause of the error. M8-robot was created by modifying a free 3D Rig for 3DStudioMax by Starostin (2007). It was selected because its face is quite expressive and the model is pre-rigged for animation. Unity3D does no support morphing, so this was the key reason for selecting this 3D model. The expressions and animations used for student feedback were created using 3DStudioMax.

Table 6.3 shows the clues and feedback provided by M8 when the student requests feedback on his/her latest error by clicking on the green question mark button in the Cockpit view. The animation of M8-robot blinking and talking is synchronised with the feedback message. However, students must attempt the challenge at least once, so M8-robot can identify appropriate feedback. When the student achieves the game/learning goals, positive reinforcement is given implicitly (using images) and explicitly (using M8-robot's response). For example, Figure 6.26 shows a screenshot captured when the student achieves the game challenge. M8-robot raises its thumb and is shown to say 'YIPEEEEE! You made it', while the player character is shown with her left fist raised in triumph. The 'Well done' text is displayed when students achieve the target within a relative error of the distance (e_d) between 2% and 5%. Table 6.4 includes the different categories that are assigned to student performance according to the score achieved, which are assigned using the evaluation performed by the cognitive model.

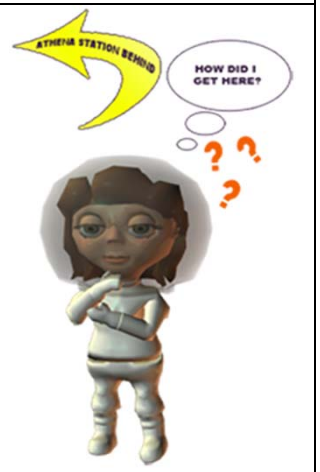
Student Misconception	*Explanation*	*Implicit feedback*	*Student Misconception*	*Explanation*	*Implicit feedback*
The student set the direction of the acceleration incorrectly.	Since the student set a positive direction for the acceleration (towards Athena), Alpha Centauri will never stop and the ship will keep going forever.		The magnitude of the acceleration is larger than $40\ m/s^2$	A person cannot bear accelerations larger than $4g$, without passing out. Where g corresponds to the value of the gravitational acceleration in the Earth ($9.81 m/s^2$).	
The time travelled (t_s) exceeded the expected time to arrive to Athena without exhausting the available fuel. i.e. ($t_s>T$).	The combustible finished before arriving to Athena, since the student did not select suitable values for the initial velocity (v_i) and acceleration (a).		Alpha Centauri passed Athena station	The student did not select a set of values for the initial velocity (v_i) and acceleration (a) that leaves Alpha Centauri a reasonable distance from Athena station (within an absolute relative distance error, $\lvert e_d \rvert$, equal to or less than 5% of D).	

Table 6.2 Implicit feedback created for PlayPhysics' game challenge

In addition, the learning companion M8-robot responds to the emotion reported by students, sometimes mirroring the emotion and in the case of negative emotions, such as sadness, trying to show affinity and by encouraging them to keep making an effort so as to achieve the game/learning goals. It is important to remember that the manner of response to emotions is a remaining challenge, since negative emotions do not necessarily have a negative influence on student learning and motivation. In this research, we are only interested in determining if students perceive the emotional behaviour of M8-robot as appropriate and if this perception is related to student gender. Table 6.5 shows the possible affective responses of the M8 robot to student self-reported affective states.

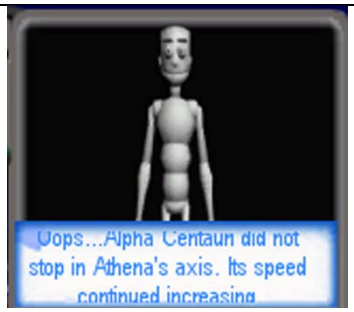
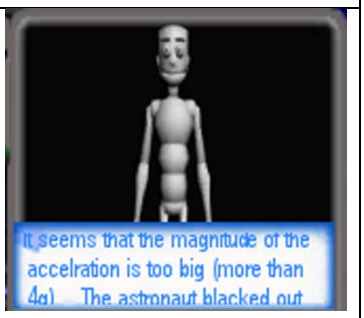
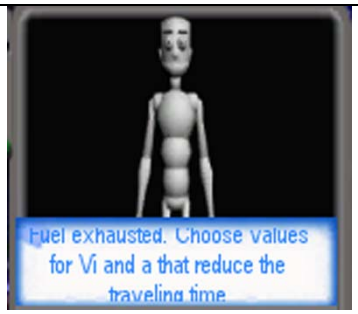
Student misconception	*Clue or explanation given by M8 robot*	*Student Misconception*	*Clue or explanation given by M8 robot*
The student incorrectly sets the direction of the acceleration.		The magnitude of the acceleration is larger than $40\ m/s^2$.	
The time travelling (t_s) exceeded the available time i.e. ($t_s>T$).		Alpha Centauri did not reach Athena.	
Alpha Centauri passed the Athena station			

Table 6.3 Clues and explanations given by M8-robot by request

In the GUIAnimation class, the UpdateAnimation function (Figure 6.27) comprises a frames array with all the textures used for the animation of one of M8-robot's behaviours. The currentFrame variable corresponds to the index of the frame that is displayed at that moment. When the currentFrame value is larger than the frames array length the currentFrame index points to the end of the animation. Therefore, this animation must be termi-

nated. The frameInterval variable is the time that must elapse before switching the frames. After starting, executing and ending any behaviour, M8-robot returns to its idle behaviour. This helps to achieve continuity and prepares M8-robot to show other behaviour in the future. The animation and message text selected depend on the type of event that was triggered by the tutor module, which may be a student misconception or self-reported emotion.

Success range	*Category granted to the student performance*
if $(mark \geq 70\%) \& (mark < 80\%)$	OK
if $(mark \geq 80\%) \& (mark < 90\%)$	WELL DONE
if $(mark \geq 90\%) \& (mark < 100\%)$	VERY WELL DONE
if $(mark = 100\%)$	PERFECT

Table 6.4 Categories of student performance according to the score achieved

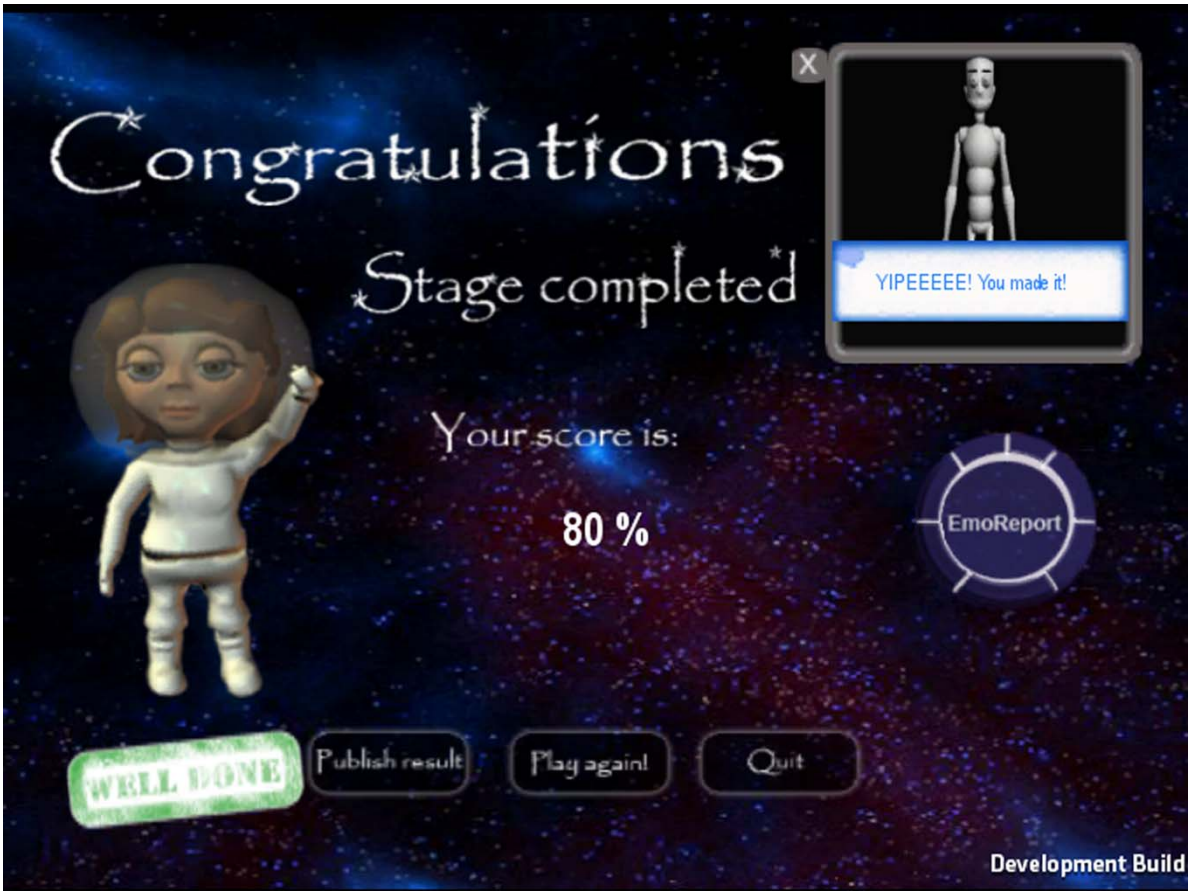


Figure 6.26 PlayPhysics acknowledging and celebrating the student success

Student self-reported emotion	*M8 affective response*	*Student self-reported emotion*	*M8 affective response*
Pride		Shame	
Joy		Sadness	
Gratitude		Anger	
Enjoyment		No emotion	
Frustration		Boredom	

Table 6.5 M8 learning companion's affective responses to students self-reported state

```
function UpdateAnimation()
{
        frameCounter -= Time.deltaTime;

         if (frameCounter <= 0)
         {
                 frameCounter = frameInterval;
                 ++currentFrame;
                 if (currentFrame >= frames.length)
                 {
                         if (AnimationEnded)
                         {
                                 animationEnabled = false;
                                 AnimationEnded();
                         }
                         currentFrame = 0;
                 }
         }

}
```

Figure 6.27 UpdateAnimation function in the Java Script class GUIAnimation

6.7 PlayPhysics implementation for the control group

As mentioned in Chapter 4, it is necessary to have a reference to determine whether students gain knowledge by interacting with PlayPhysics' GBL environment. Hence, it is necessary to establish what average student performance is like without using PlayPhysics' GBL environment. This is obtained by using the control group. Alternative functionality was created in PlayPhysics for students in this group, where they can access a PowerPoint presentation in the form of a video demonstrating the same concepts explored by PlayPhysics and reviewing solved examples (Figure 6.28). This same presentation can be accessed by students belonging to the focus group that also interact with PlayPhysics GBL environment when they click on the glass magnifier next to M8-robot (Figure 6.29). The presentation opens in a pop-up window. Students in the focus and control groups have access to the same pre-test and post-test (Appendix H). However, only students in the control group have access to the qualitative evaluation of PlayPhysics GBL environment.

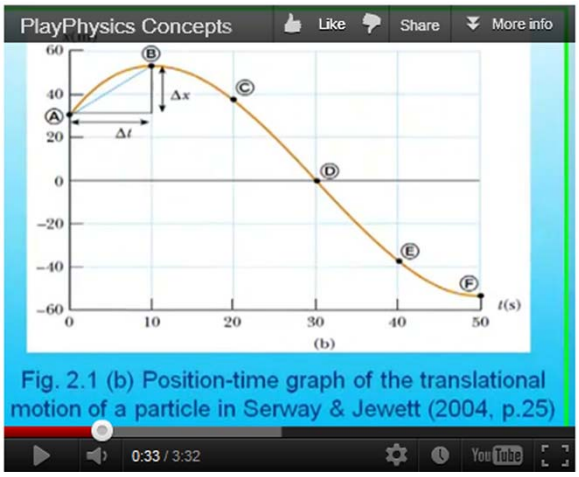


Figure 6.28 Screenshot of PlayPhysics' PowerPoint presentation video

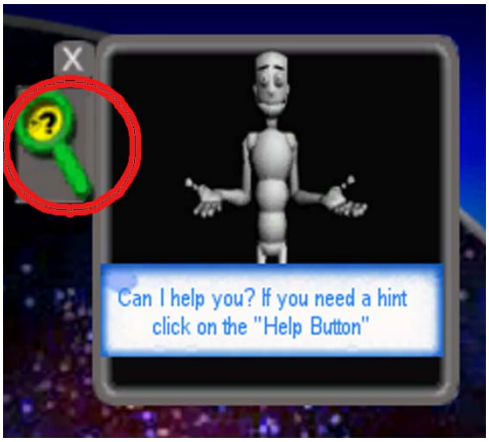


Figure 6.29 Accessing the PowerPoint presentation of physics concepts

6.8 PlayPhysics interaction and execution examples

This section shows some of the interaction examples of PlayPhysics involving the Head of Department, System Administrator and Students in order to provide an extended illustration of its capabilities and available functionality.

6.8.1 Student registers in PlayPhysics

Lecturers assign students to Control and Focus groups based on their past performance in the module, ensuring that the groups are balanced. Once students know their assignation, they can register in PlayPhysics by accessing the link http://elearning2.ccm.itesm.mx:8281/PlayPhysics/html/index.htm as 'Students'. Then the student clicks on 'Register' (Figure 6.30) in order to select their university, gender and investigation group (Figure 6.31). The latter can be Group A or B, where 'Group A' corresponds to the Control group and 'Group B' corresponds to the Focus group. Figure 6.32 shows how students input their registration number, personal details and password and choose their group

before finally clicking the Register button. A message saying that the registration was successful is displayed on the screen.

Figure 6.30 The student chooses to register in PlayPhysics

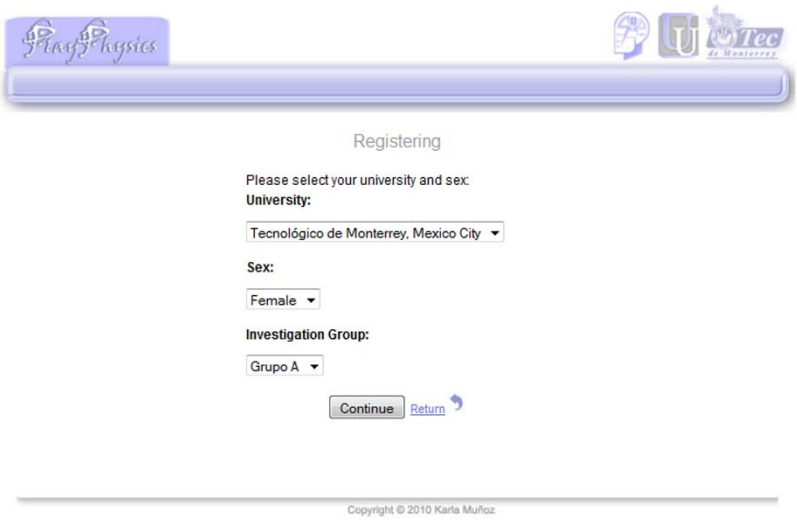


Figure 6.31 Students choose options to register in PlayPhysics

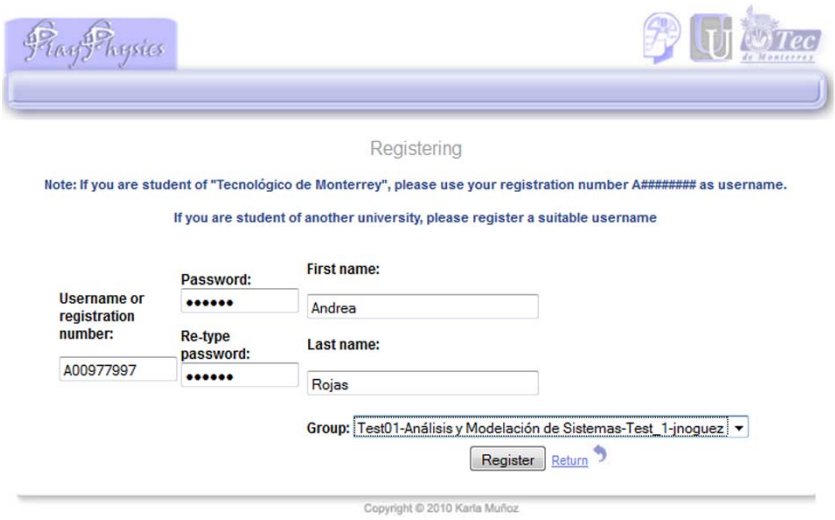


Figure 6.32 Students introducing personal details and choosing their group

6.8.2 Control group student interacting with PlayPhysics

To interact with PlayPhysics, students in the Control group input their username and password to log in (Figure 6.33). Then in their work tray they can see the 'do pre-test' link, (Figure 6.34). If students click on it, they are taken to the 'Pre-test'. This is displayed as shown in Figure 6.35. They can submit the test by clicking on the Submit button. After answering the pre-test a message will appear on the screen saying 'Thank you for answering the pre-test', which is followed by displaying the mark obtained. This stage of answering the Pre-test is the same for students in the Focus group. When students return to the work tray, they can see the mark obtained on the pre-test and the active link, entitled 'Physics concepts' (Figure 6.36).

Figure 6.33 Students in the control group logging in

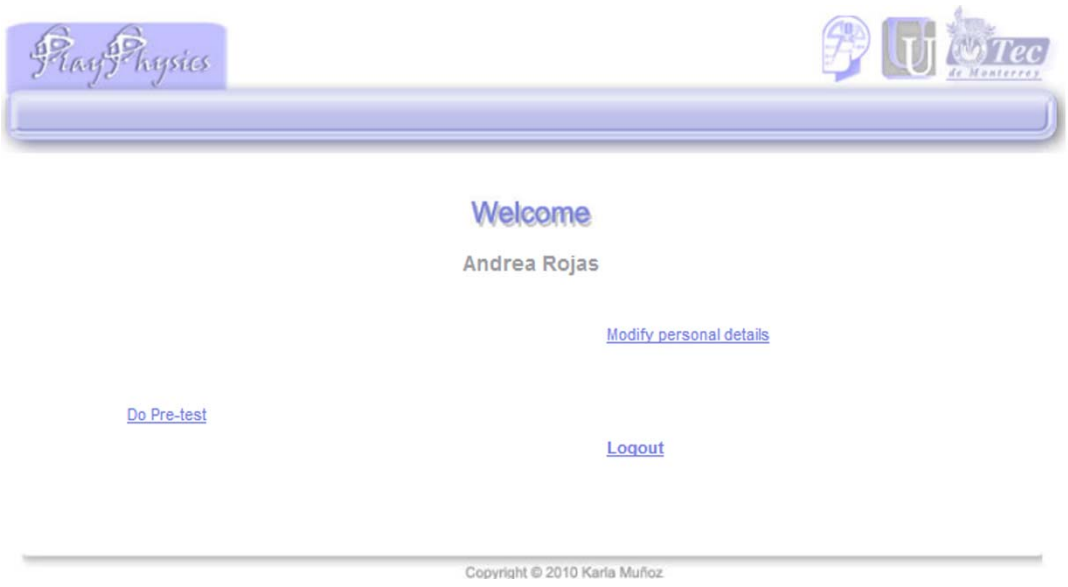


Figure 6.34 Initial state of the work tray for students in the control group

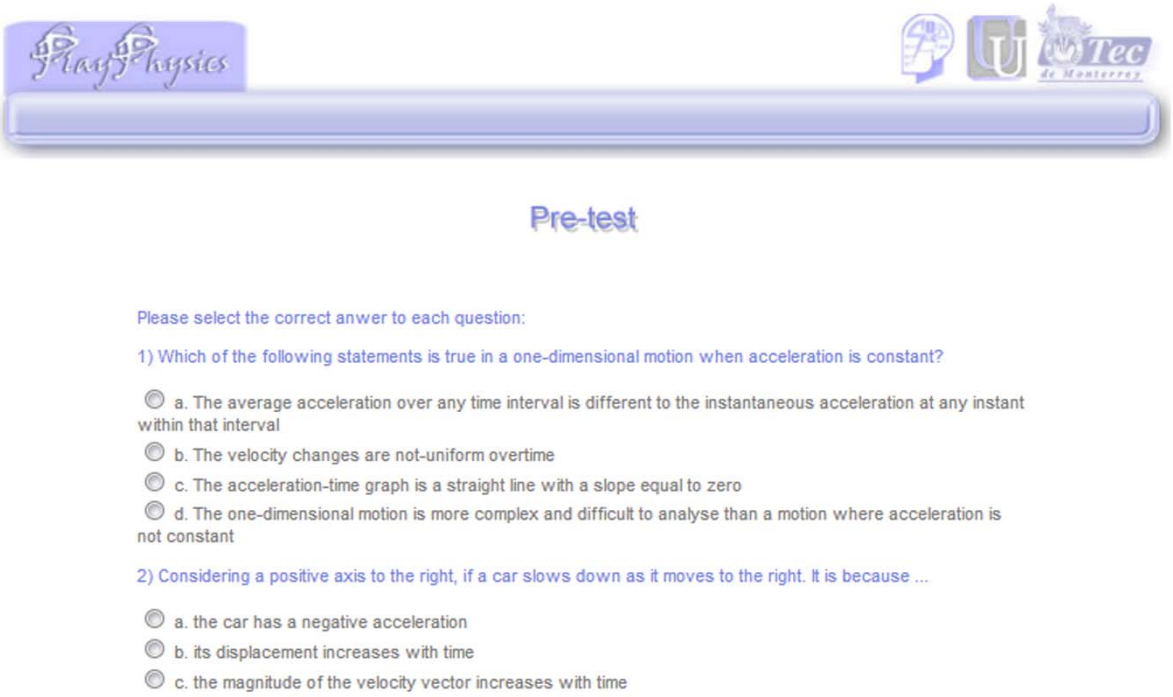


Figure 6.35 Pre-test displayed on the screen to be answered by students

By clicking the link 'Physics Concepts', a video about physics concepts exemplifying how related exercises are solved is played (Figure 6.36). Students can watch the video at their own pace and review it as many times as they wish before proceeding to solve the Post-test. Once students proceed to solve the Post-test the presentation video is not longer available. The only available link now in the student work tray is 'Do Post-test' (Figure 6.37).

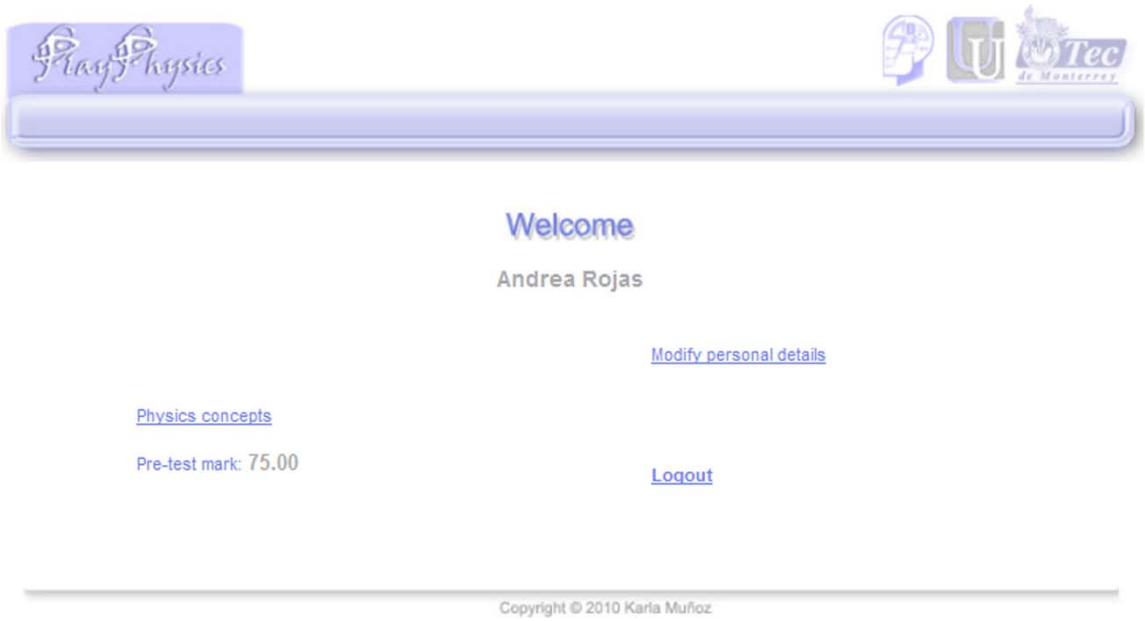


Figure 6.36 Work tray showing the active link physics concepts

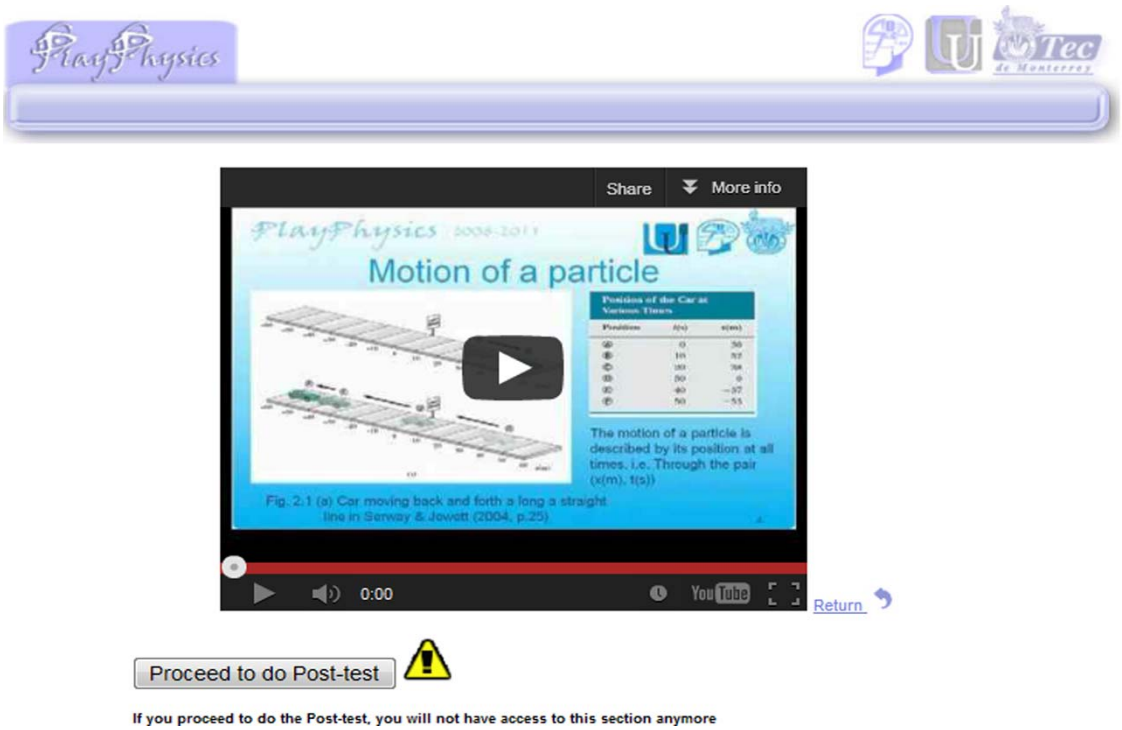


Figure 6.37 Video of the PowerPoint presentation with physics concepts

Later, students can access the Post-test link see Figure 6.38. At this stage of interaction, the functionality offered by PlayPhysics is the same for students in both groups: Control and Focus. The Pre-test and the Post-test can be solved only once (Figure 6.39). After answering the post-test, students submit it and receive a message saying: 'Thank you for answering this test', which is accompanied by the mark obtained. The mark, as on the pre-test, can be a value between 0% and 100% and it is also displayed in the work tray (Figure 6.40). Students can also modify their personal details, change their password or log out.

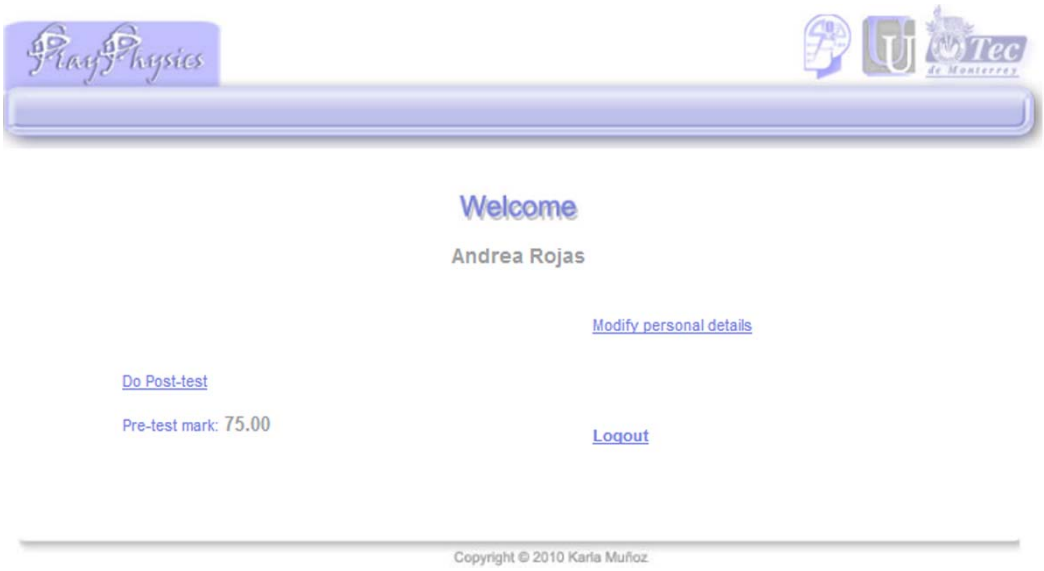


Figure 6.38 Student work tray showing that the Post-test link is available

Figure 6.39 The Pre-test is displayed to students

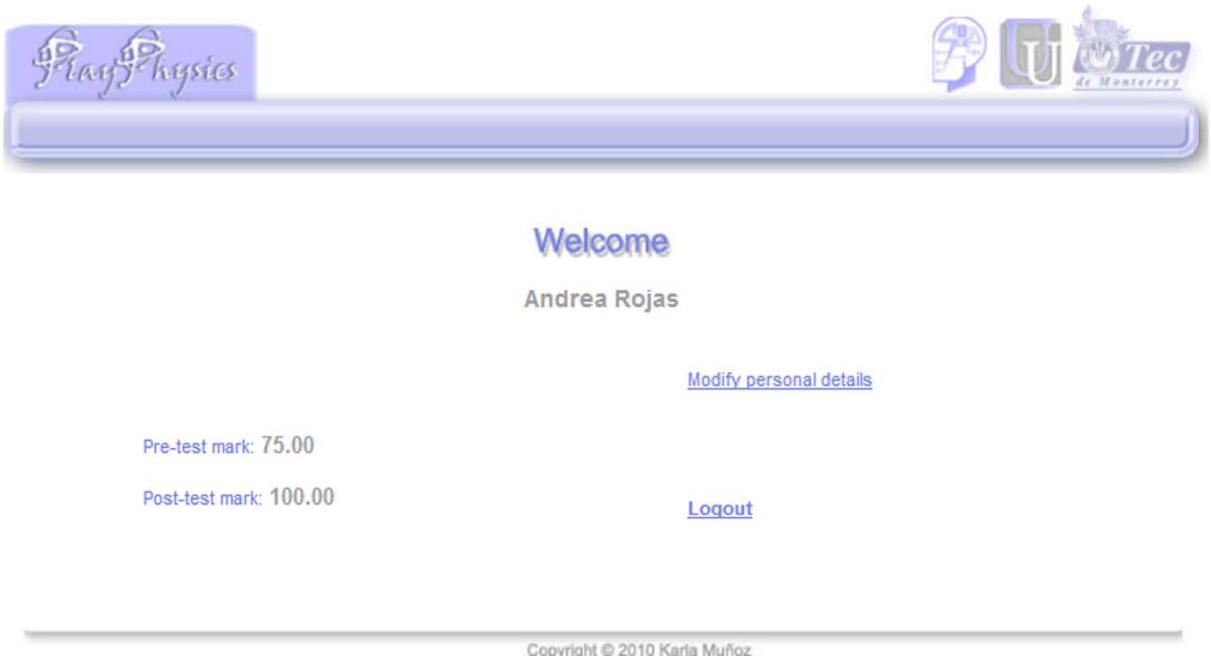


Figure 6.40 Students can consult the marks of both tests in his work tray

6.8.3 Student interacting with PlayPhysics educational game

Students in the focus group have to solve the pre-test in order to access to the PlayPhysics educational game. When the link 'Launch PlayPhysics' is displayed (Figure 6.41) students can click on it to start to interact with the PlayPhysics educational game (Figure 6.42).

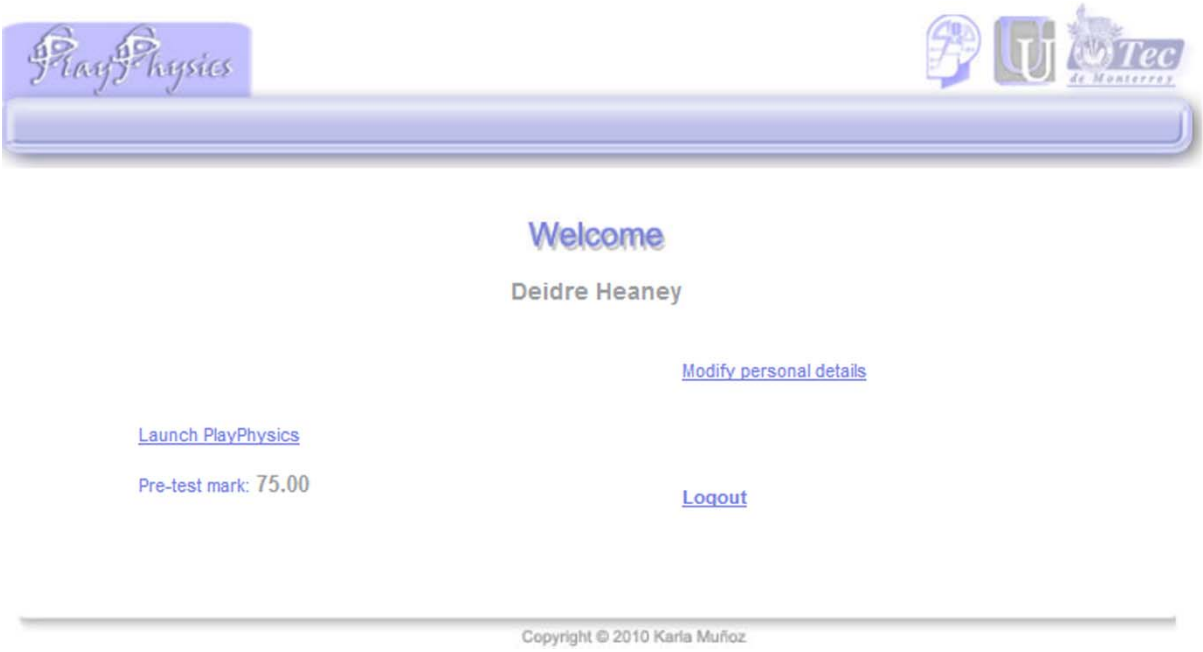


Figure 6.41 Work tray of students in the focus group ready to start playing

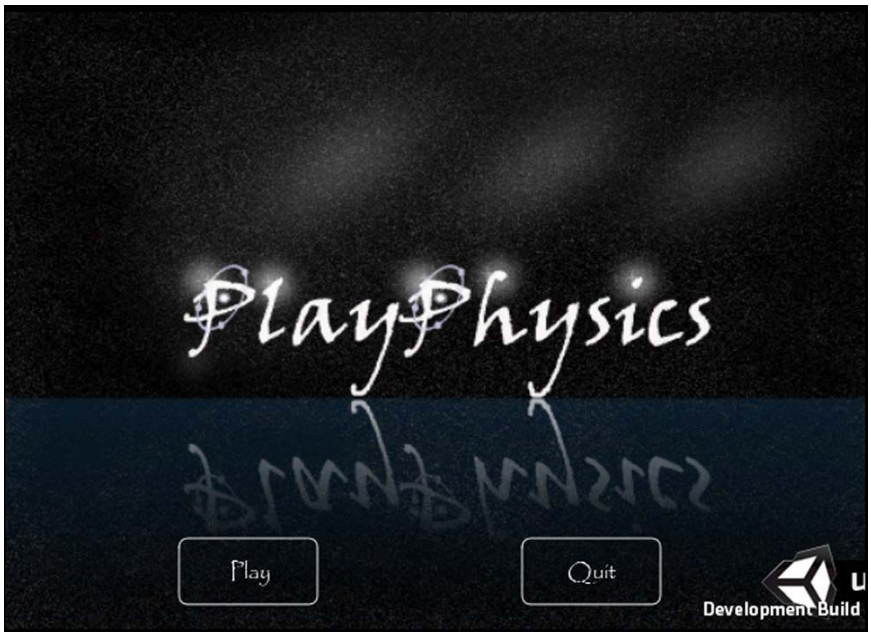


Figure 6.42 Screenshot showing the main page of PlayPhysics educational game

In the game, the story is introduced through a series of cut-scenes (Figure 6.43), the pace of which is driven by students. The game challenges are then introduced by M8-robot and a brief tutorial related to the functionality of the game challenge GUI is also provided (see Appendix G). In Figure 6.44 M8-robot explains to students how to achieve the game challenge goal, while Figure 6.45 shows students how to use the game challenge GUI in order to observe their progress. A key point of this tutorial is that M8-robot explains to students how to report their emotion (Figure 6.46).

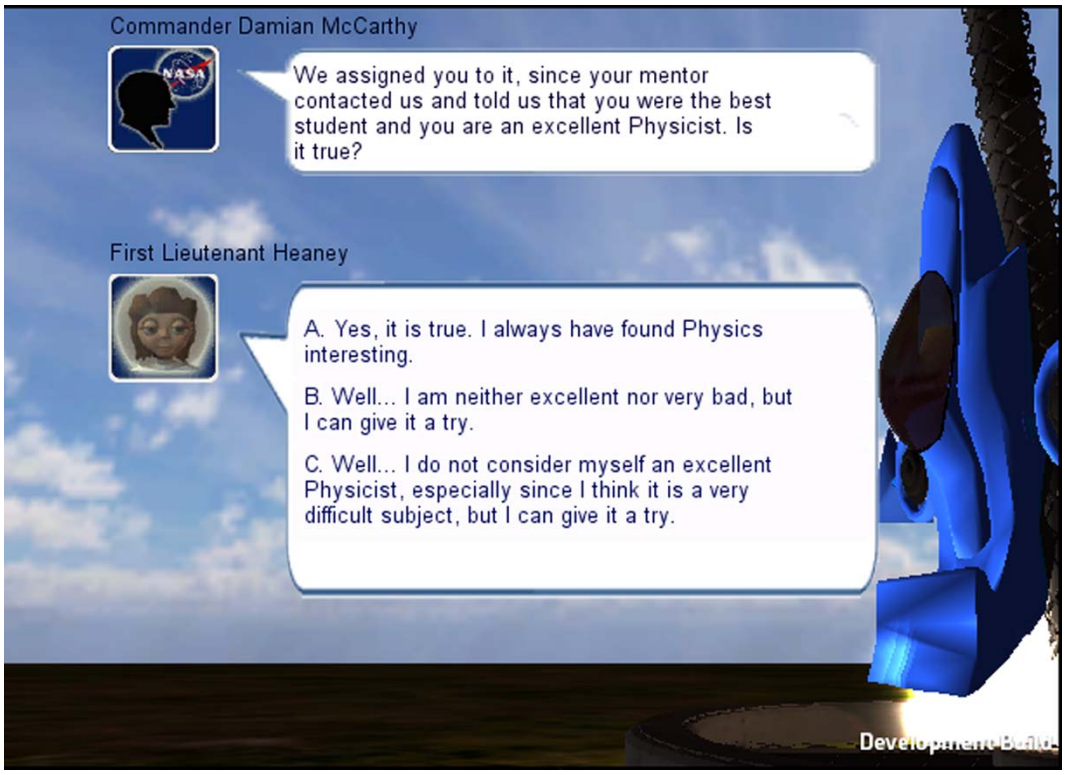


Figure 6.43 Screenshot showing a cut scene of PlayPhysics educational game

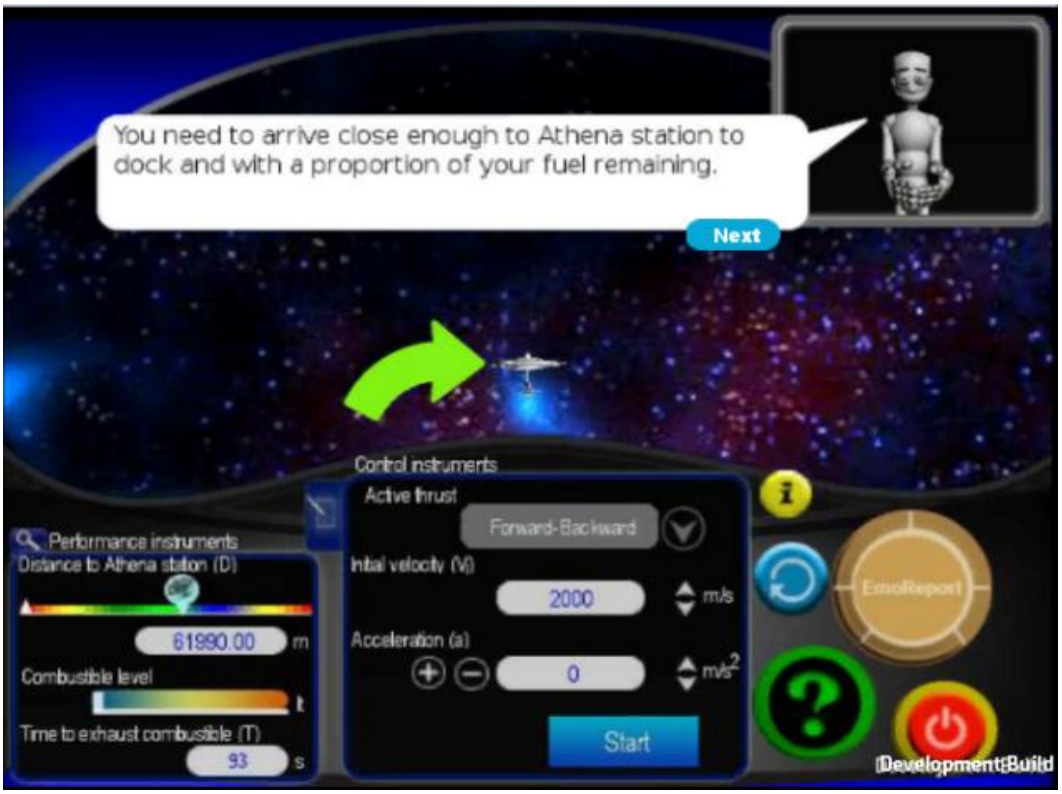


Figure 6.44 M8-robot Explaining PlayPhysics game challenge goals

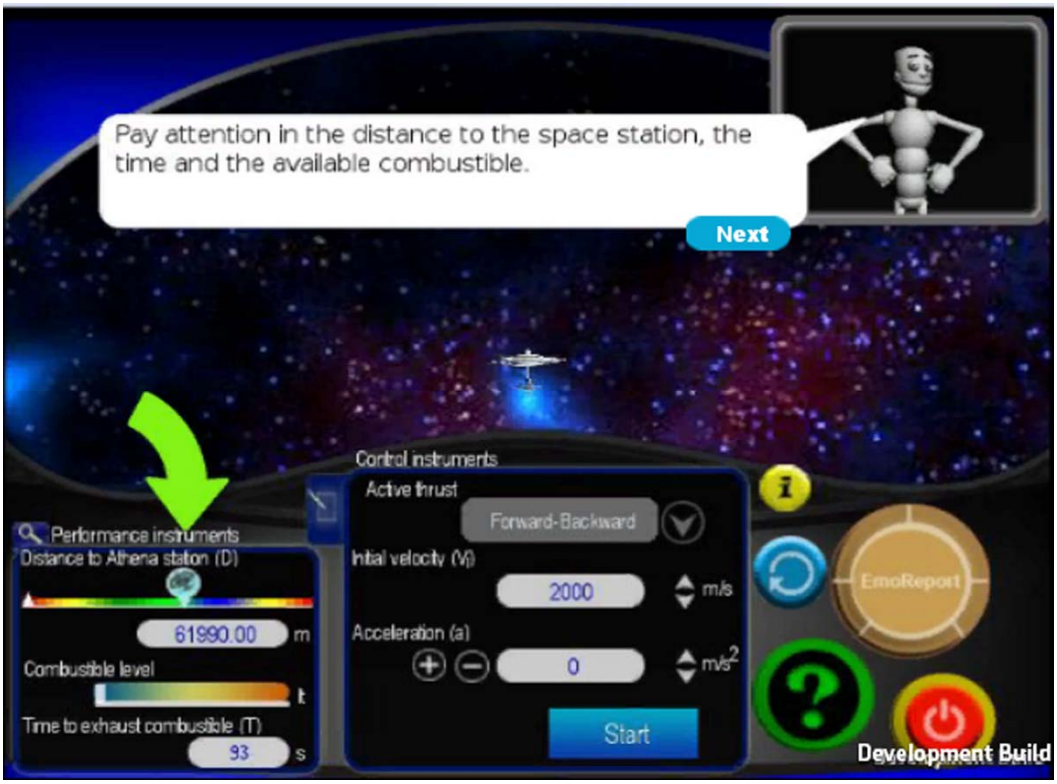


Figure 6.45 M8-robot explaining the functionality in the game challenge GUI

Figure 6.46 M8-robot briefly explains how students can report their emotion

At this point, students can start to interact with PlayPhysics' game challenge. In this case, PlayPhysics randomly assigns the key variables: a) distance to Athena station and b) time to

exhaust the combustible: $D = 30240\ m$ and $T = 86\ s$. Students set Alpha Centauri's initial velocity as $v_i = 1688\ m/s$ and it's acceleration as $a = -33\ m/s^2$. Users then click the Start button. M8-robot periodically asks students to report their emotional state (Figure 6.47).

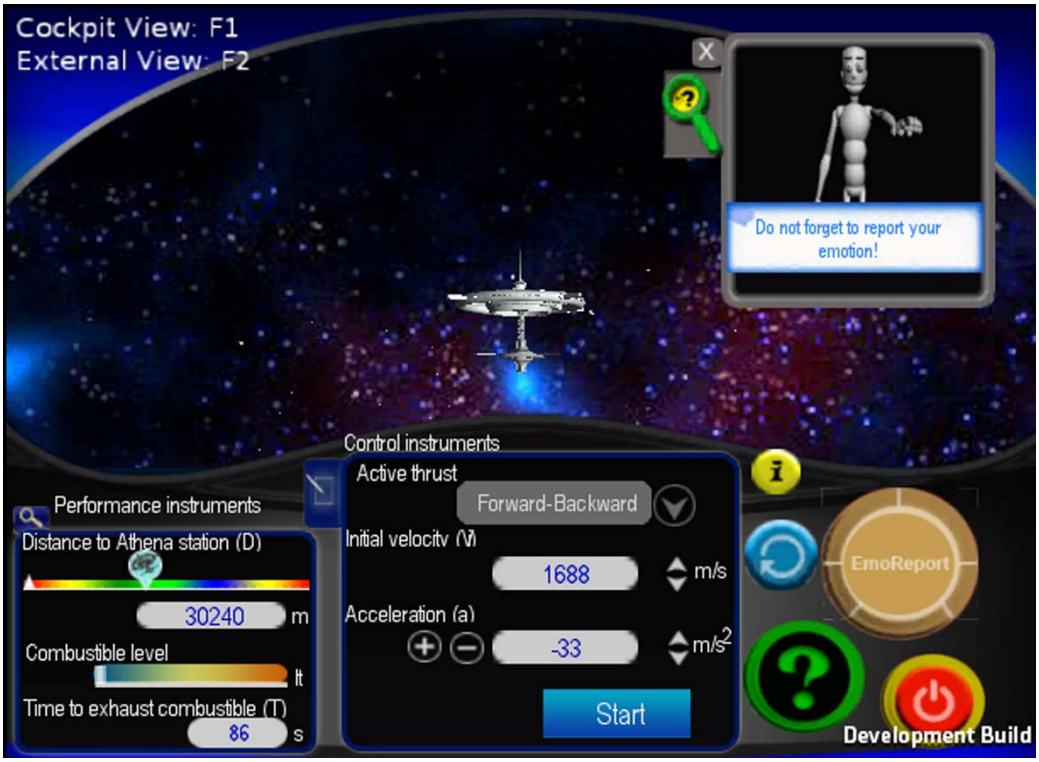


Figure 6.47 Students solving the challenge and M8-robot requesting emotional input

PlayPhysics reviews the direction of the acceleration, the magnitude of the acceleration and calculates d_s and t_s in order to evaluate student choices and identify any misconception. In this case, the direction of the acceleration is negative, indicating a deceleration, which reflects that the student knows to apply a deceleration in order to stop (this and other assumptions are derived from the knowledge provided by the domain expert. See Chapter 5, section 5.3.1). The magnitude of the acceleration is then checked to ensure it is not over $40\ m/s^2$, since a person cannot tolerate acceleration over $4g$ without losing consciousness. The magnitude set by the student in this case is $33\ m/s$. Therefore, PlayPhysics assumes that the student is aware of this fact and proceeds to calculate the values for $d_s = 43172\ m$ and $t_s = 51\ s$. Since $t_s<T$, the distance needs to be compared. This is achieved by calculating the relative error (e_d), in this case $e_d = 23.48\% \approx 24\%$. Since this error is larger than the allowed, i.e. ±10%.

The student did not succeed in the game challenge and a message saying 'You Died' is shown on the screen (Figure 6.48). Students should observe that they overshot Athena and that Alpha Centauri stops some distance beyond the target. Students are then presented with their score and implicit feedback (Figure 6.49). In this case, the player character is depicted thinking 'How did I get here' and an arrow shows that Athena station is some distance

away. Appropriate music is played. Students must report their emotion to M8-robot after receiving the outcome. Students then have two options: play again or quit.

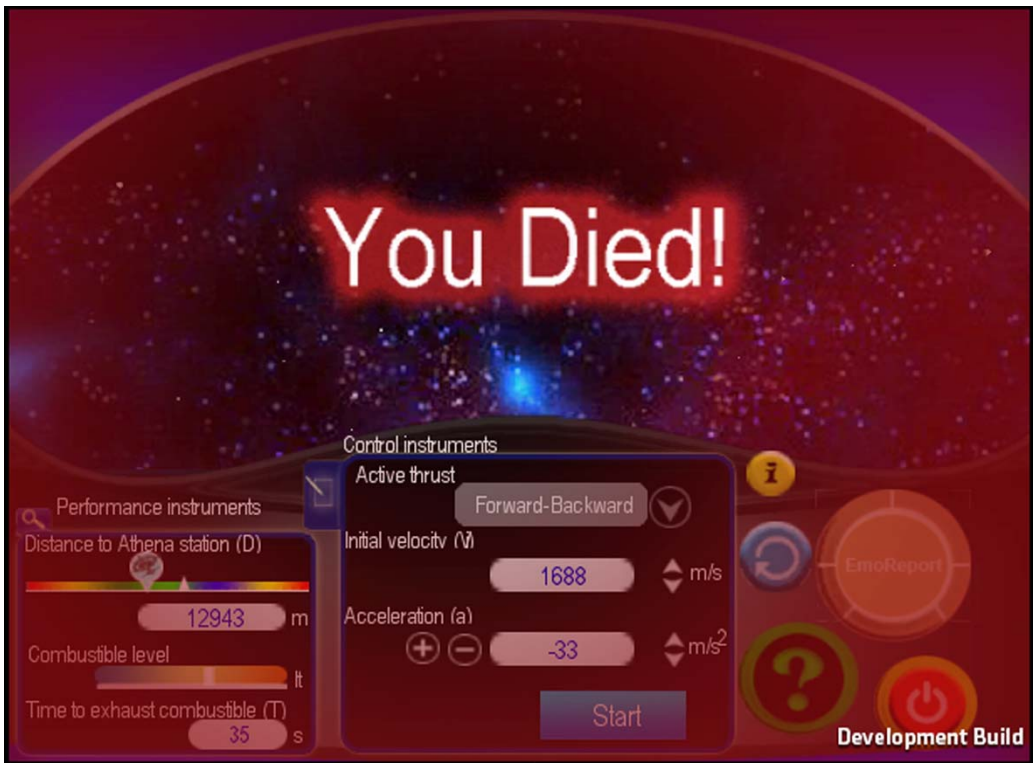


Figure 6.48 PlayPhysics communicating to the student that he/she failed

Figure 6.49 Students in the focus group receiving feedback

Consider the case of a male student who is presented with the restriction variables $D = 19470\ m$ and $T = 102\ s$. After analyzing the problem, he selects and sets the values for v_i and a as $v_i = 1242\ m/s$ and $a = -39\ m/s^2$. It can be observed that the direction and magnitude of the acceleration indicates that the student does not have any problem with the concepts related to these aspects of the acceleration. Therefore, PlayPhysics calculates d_s and t_s, where $ds = 19776\ m$ and $t_s = 32\ s$. Since $t_s<T$, the relative error of the distance is obtained, where $e_d = 1.57\%$.The relative error is between 2% and -2%, as a result, it is necessary to compare the time calculating t_{max}, t_{min} and Δt to determine the accuracy of the solution. In this case $t_{min} = 31\ s$ $t_{max} = 26\ s$ *and* $|\Delta t| = 5\ s$. In this case, t_s is in the range between $[t_{min}, t_{min} + |\Delta t|/3]$ and a perfect score has been achieved.

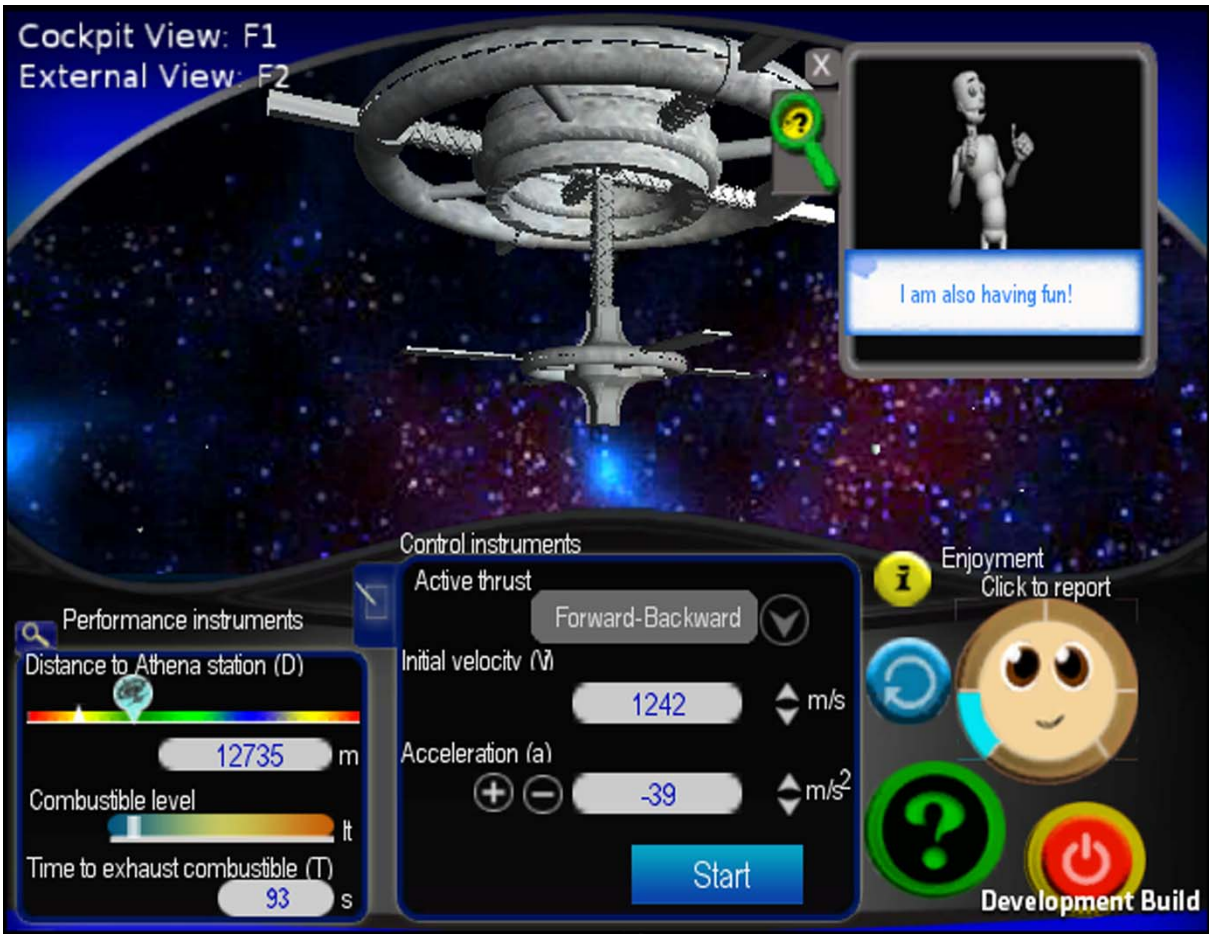


Figure 6.50 Student interacting with PlayPhysics game challenge

The student sees how Alpha Centauri approaches Athena and comes to rest beneath it (Figure 6.50). While interacting with PlayPhysics, the selected contextual variables, e.g. mouse position (*mouse_focused_coarse_value*) and the number of times help was required (*num_times_help_asked*), are monitored and recorded in the *stdobj* object instance of the *StudentData* class. If the GSR signal of this student is being recorded, timestamps of the GSR signal can be matched with timestamps corresponding to the contextual and interaction

variables in order to have a broad perspective to reason about student emotion. Figure 6.51 shows a fragment of code in StudentData corresponding to the *sendDuringInteractionData* function, which shows all the variables which are monitored and sent through POST to be saved in PlayPhysics' database. In addition, the student may report his emotion during the PlayPhysics game challenge. Finally implicit feedback is given in conjunction to a score of 100% and congratulatory music is played (Figure 6.52). The student is asked to report his emotion regarding the score achieved. At the end the student is also presented with the options: play again, quit and publish result. If the option quit is selected, the student can see the top 10 scores before returning to the work tray.

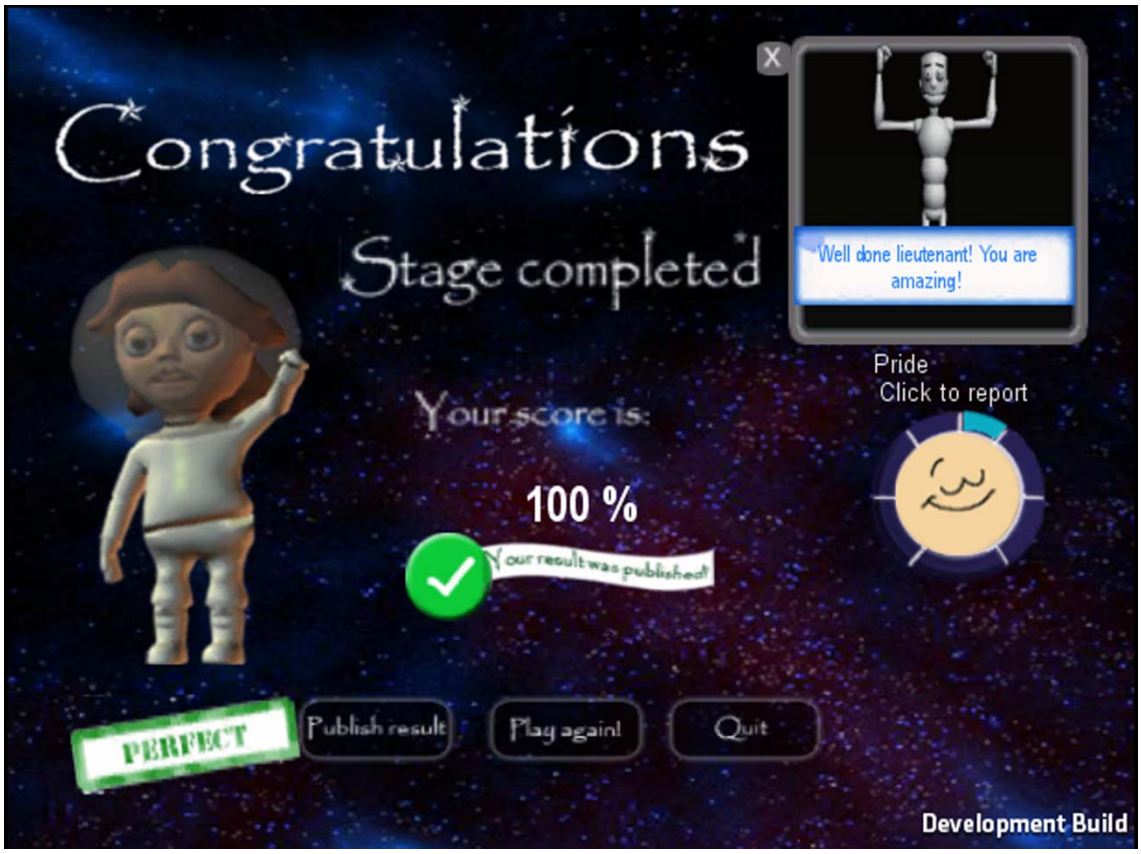


Figure 6.51 The student is congratulated for his score and he reports his emotion

```
function sendDuringInteractionData()
{
        ...

        form = new WWWForm();
        form.AddField("username_regisnum",GetStudentID());
        form.AddField("time_started",GetTimeStarted());
        form.AddField("outcome_percentage", GetOutcomePercentage());
        form.AddField("outcome", GetOutcome());
        form.AddField("id_outcome", GetEndCondition());
        form.AddField("num_times_help_asked", GetNumTimesHelpAsked());
        form.AddField("num_times_help_received", GetNumTimesHelpReceived());
        form.AddField("num_attempts_alone", GetNumAttemptsAlone());
        form.AddField("independence", GetIndependence());
```

```
            form.AddField("total_attempts", GetTotalAttempts());
            form.AddField("quality_of_tutor_feedback", GetQualityOfTutorFeedback());
            form.AddField("average_tutor_feedback", GetAverageTutorFeedback().ToString("f2"));
            form.AddField("interaction_interval", GetInteractionInterval());
            form.AddField("interaction_interval_seconds", GetInteractionIntervalSeconds());
            form.AddField("time_ended", GetTimeEnded());
            form.AddField("mouse_focused", GetMouseFocused());
            form.AddField("mouse_focused_coarse_value", GetMouseFocusedCoarsedValue().ToString("f2"));
            form.AddField("time_to_achieve_goal", GetTimeToAchieveGoal());
            form.AddField("result_published", GetResultPublished());
            form.AddField("emotion_reported", GetEmotionReported());
...
}
```

Figure 6.52 SendDuringInteractionData function in StudentData

6.8.4 Head of Department registering lecturers and groups

Once the Head of Department is registered in PlayPhysics by the System administrator, he/she can login by accessing the URL: http://elearning2.ccm.itesm.mx:8281/PlayPhysics/. (Figure 6.53). Users can then click on the Staff button and access to the login page (Figure 6.54). At this point, the Head of Department will have access to his/her work tray (Figure 6.55) where the 'Maintenance of Lecturers' option can be selected in order to create, delete or modify lecturers. In this case, the Head of Department introduces all the information related to the lecturer and clicks the Register button (Figure 6.56). A message should then appear on the screen confirming that the registration in PlayPhysics was successful.

Figure 6.53 Accessing PlayPhysics' main page and login as a Staff member

If the Head of Department wants to register a group, he/she selects the link for the 'Maintenance of Groups' option in the work tray. Figure 6.57 shows how to register the group.The user inputs the name of the group and selects the corresponding semester, subject and lecturer. It is important to remember that the semester and the subject are registered earlier in the PlayPhysics' database by the System Administrator. Finally the Head of Department clicks the Register button. If the database operation was successful a success message is displayed.

Figure 6.54 Head of Department logging in

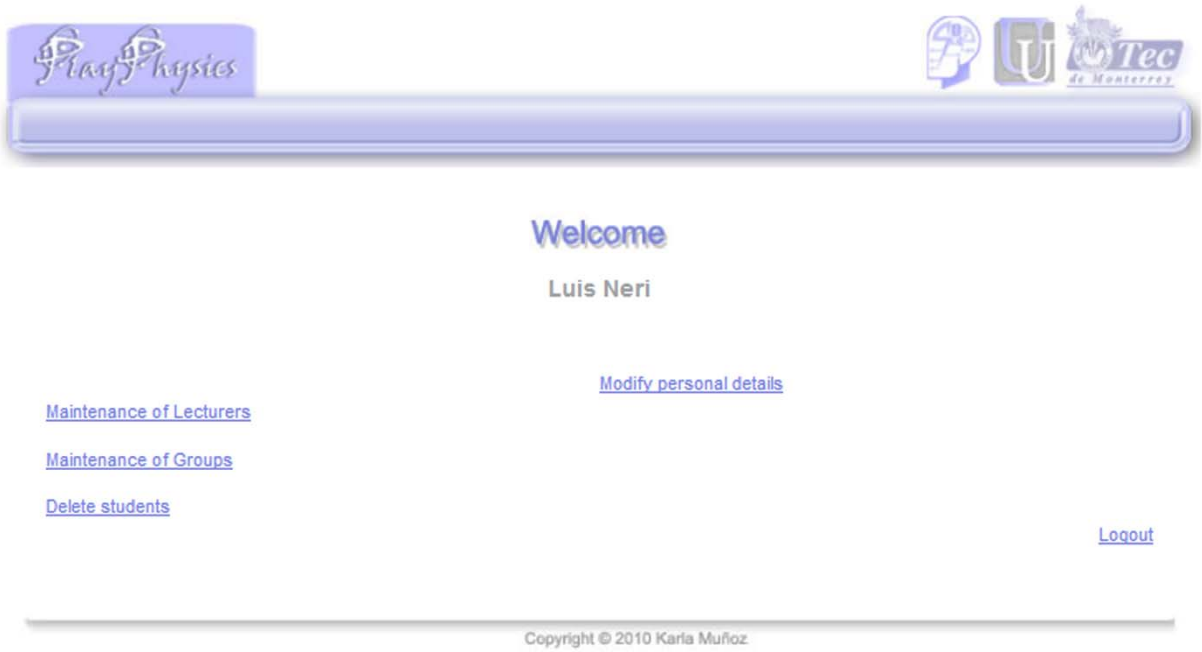


Figure 6.55 Head of Department work tray

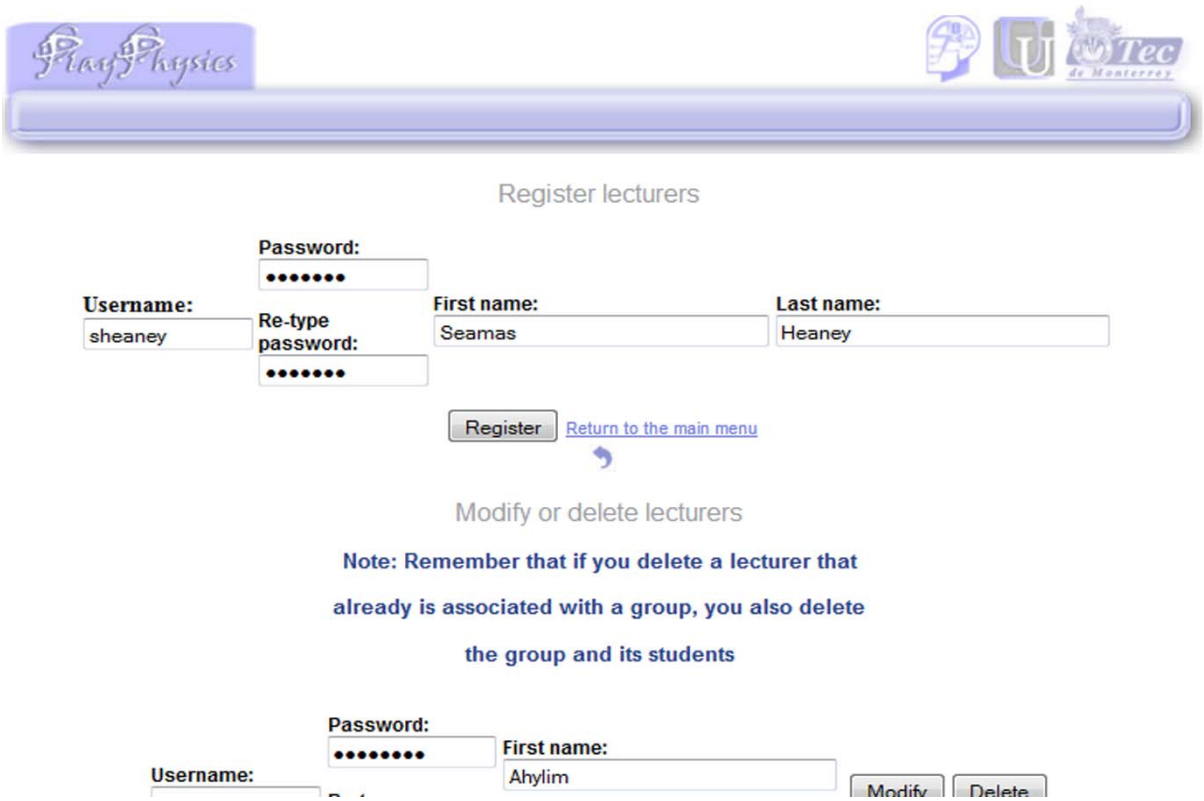


Figure 6.56 The Head of Department fill in the details of the lecturer

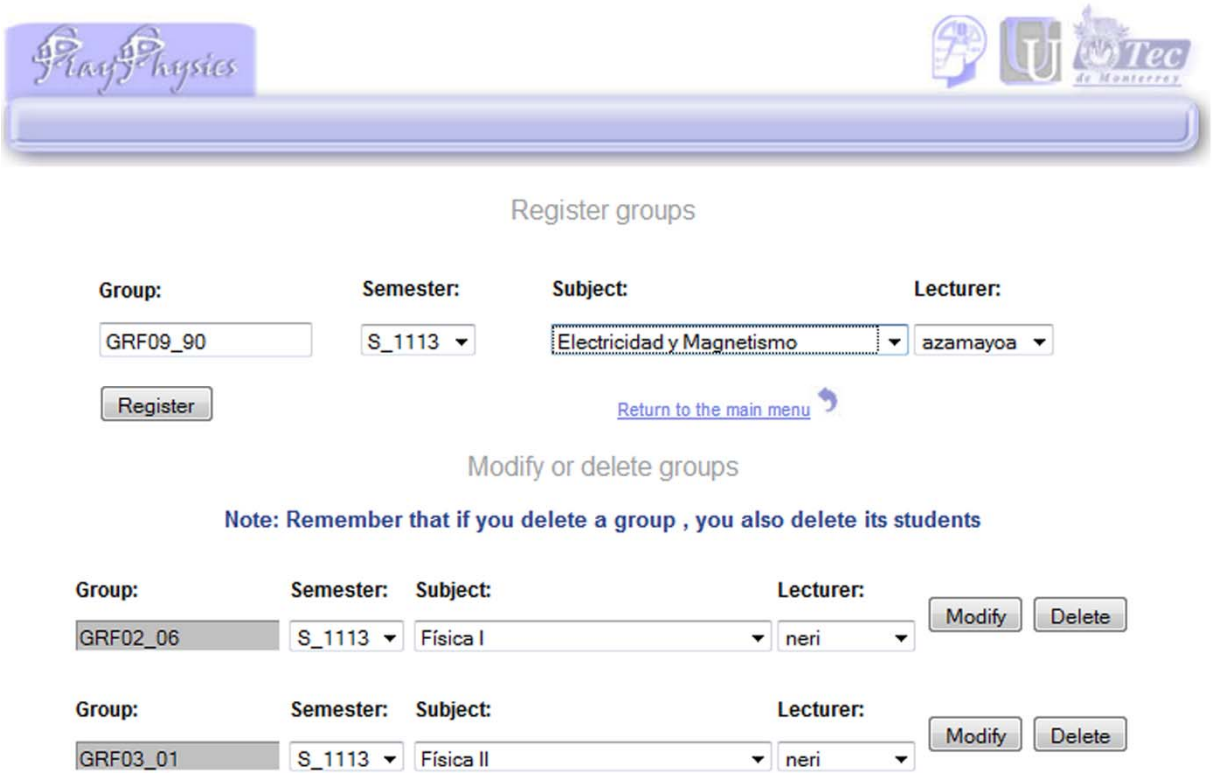


Figure 6.57 Example of a Head of Department registering a group in PlayPhysics

6.8.5 System Administrator records/monitors a student GSR signal

The System Administrator is the only person that can connect or disconnect a Bluetooth GSR sensor to students. To achieve this goal the System Administrator logs in and selects the 'Connect GSR sensor to one student' option (Figure 6.58).

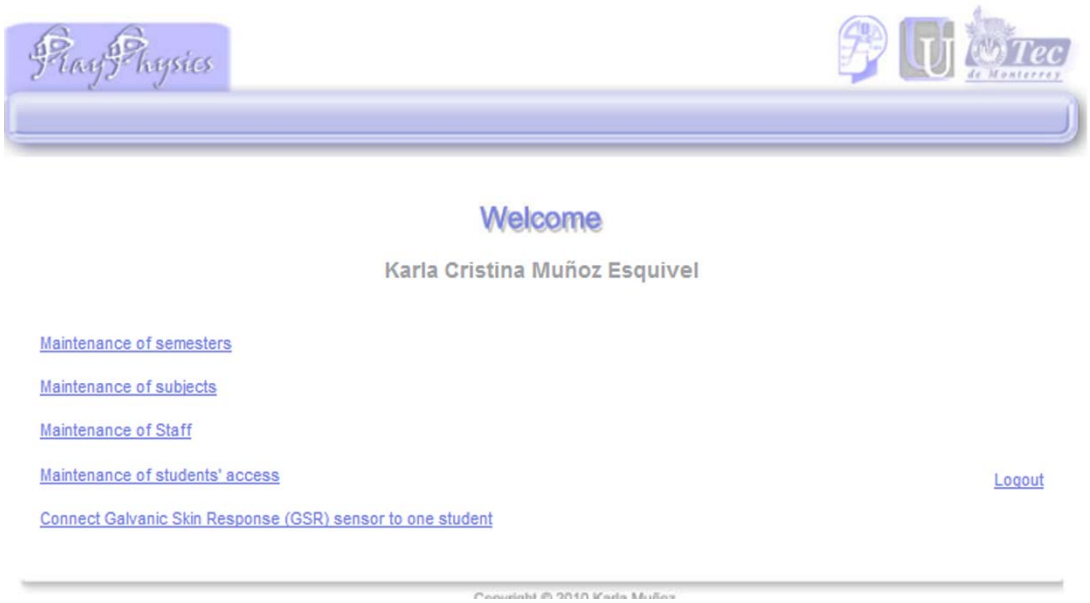


Figure 6.58 View of the System Administrator work tray

The System Administrator then selects the student group and then selects the desired student and inputs the name of the NXT brick to which the GSR sensor is attached (Figures 6.59 and 6.60).

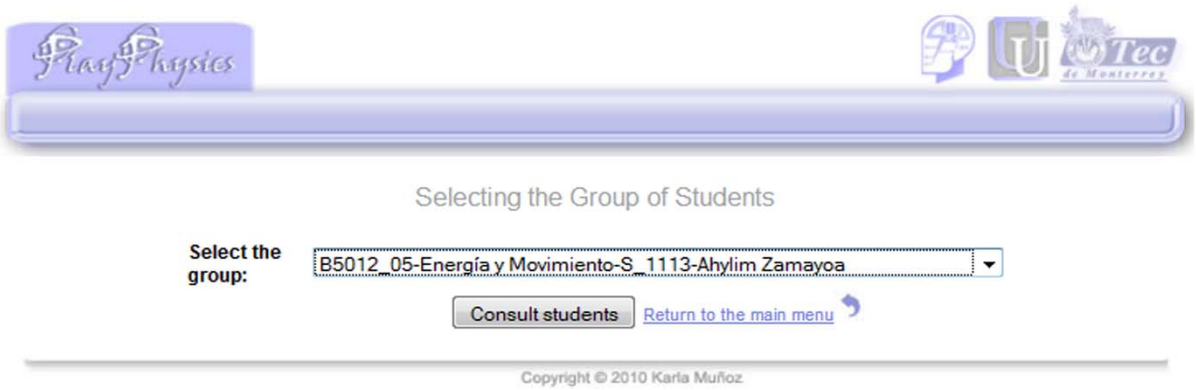


Figure 6.59 The System Administrator selects the group of the student

Once the sensor is successfully connected the control interface is shown, which allows the System Administrator to end the NXT connection (Figure 6.61). The System Administrator must ensure that the batteries of the NXT brick are charged and that the sensor shows a value other than 1023 (an indicator that the sensor is not properly attached and that the circuit is open). Another reason to ensure the batteries have a good charge not to bias the

measurement procedure. In addition, since a small current is passed through students fingers with the risk of causing a mild rash, the *Pre-tens* skin preparation gel by Tyco Healthcare is applied to limit this effect. However, this gel also affects the GSR measurements taken, so to homologate the measurement procedure, the System Administrator must ensure that the gel is applied to each student.

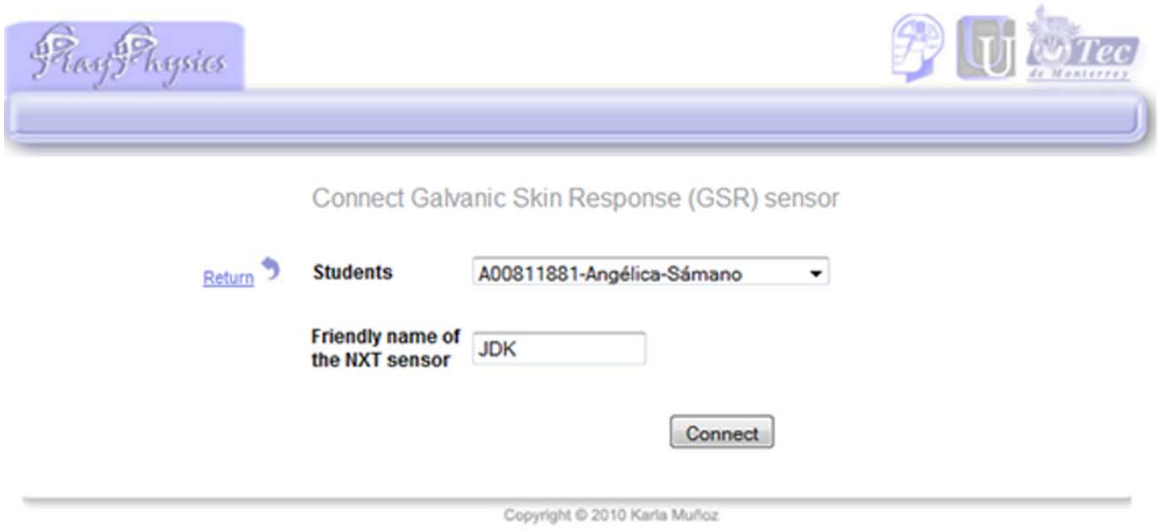


Figure 6.60 The System Administrator connecting the student with the NXT brick

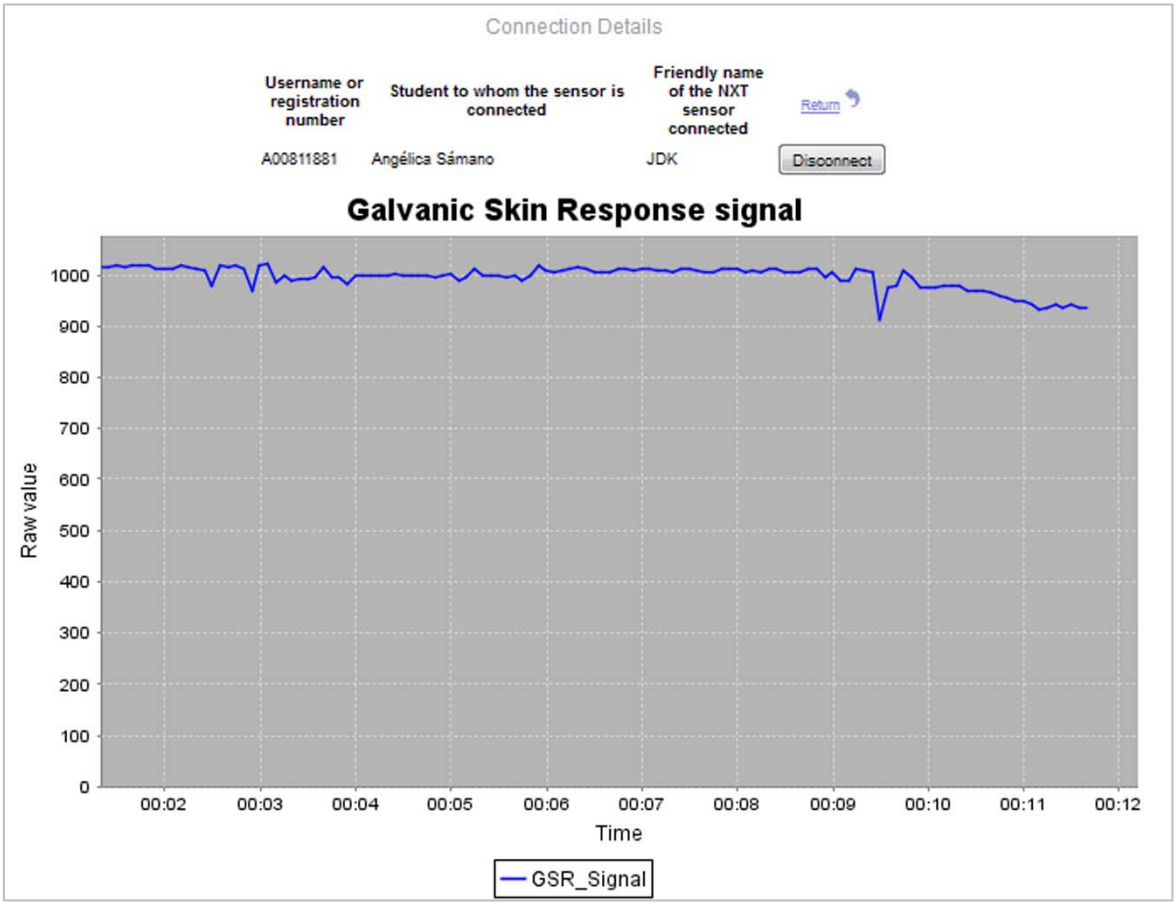


Figure 6.61 GSR signal of the student visualised by the system administrator

6.9 Summary

This chapter discussed the implementation of PlayPhysics. The manner in which Olympia's modules were implemented in the PlayPhysics' case study was shown in conjunction with the manner in which the random variables of PlayPhysics' emotional student model are acquired. Therefore, key points addressed during PlayPhysics' implementation were the acquisition and recording of input data related to the context of student interaction, behaviour and emotional state. In addition, the problems of acquiring data from student bio-signals and self-reporting were explored. Self reporting was achieved through the EmoReport wheel, which enables students to report their emotion voluntarily, i.e. whenever they wish, or at the request of M8-robot. An XML format for game dialogues was created in order to facilitate the creation of cut scenes. In PlayPhysics cut-scenes are employed to interrogate student beliefs related to students' emotional states.

The acquisition of GSR bio-signals was achieved through the creation of a GSR Bluetooth sensor, which in PlayPhysics is currently limited to reading data from one student at a time. Whilst this research assesses how bio-signal variables perform in comparison with contextual variables when reasoning about student emotion, it is a secondary goal. PlayPhysics cognitive student model is implemented as a series of production rules. Implicit feedback in the form of graphics were also created and accompanied by music that aims to encourage students. Determining the appropriate answer to students' emotional states is an ongoing challenge. As a result, M8-robot is limited to mirroring student self-reported behaviour in order to show affinity, the goal being to discover if the behaviour of M8 is considered appropriate from the student viewpoint.

Examples were given to illustrate PlayPhysics available functionality for different users, Student, System Administrator and Head of Department. These examples explained how students in the control and focus groups interact with PlayPhysics, how to connect a GSR sensor to a student and how to monitor the signal. Olympia is comprised of several functional modules, which allow the capturing of the required contextual or biofeedback variables to derive and evaluate the proposed affective student model. The following chapter explains how this model was derived and assessed though statistical methods. In addition, the qualitative and quantitative assessment of PlayPhysics as a pedagogical and affective tool is conducted. The objective is to acquire further insight into student achievement emotions in GBL environments.

Chapter 7: Experimental Results

This chapter discusses the experimental results of testing our computational emotional student model within PlayPhysics and the thesis hypothesis, i.e. the model will achieve a reasonable accuracy of classification of student emotions in GBL settings (not random). A dynamic sequence of BBNs is employed to create this model. There are two stages involved in the creation of Bayesian networks: (1) learning the network structure and (2) learning the network parameters. To learn the network structure, we employ the 'Necessary-Path Condition' algorithm in combination with the information obtained from Pearson correlations and the results of applying Binary or Multinomial Logistic Regression (MLR), depending on what best suits the observed cases. The main objective is to acquire meaningful information to solve the uncertain relations. For parameter learning, we employ the Expectation Maximisation (EM) learning algorithm. Additionally, the performance of the derived emotional student model using only contextual variables with the emotional student model including physiological variables, e.g. Galvanic Skin Response (GSR) signals, is evaluated. The purpose of these experiments is to know whether the proposed emotional student model can reason about student emotions in GBL environments with reasonable accuracy. PlayPhysics is also assessed qualitatively and quantitatively in order to achieve more insight about potential supplementary factors influencing student achievement emotions in GBL environments. Finally, an overall discussion of the results and the experience of achievement emotions in PlayPhysics is given.

7.1 Experiments design: participants, samples and tests

This section discusses the design details of the experiments that were implemented to assess PlayPhysics' emotional student model with and without physiological signals and PlayPhysics GBL environment.

7.1.1 Participants

For testing PlayPhysics, instead of approaching students directly, it was decided to approach lecturers who were teaching an Introductory Physics course at Tecnológico de Monterrey, Mexico City Campus (ITESM-CCM). This decision is due to the requirements of this institution to conduct research involving human participants and the requirement that lecturers

make students aware of the goals of this investigation and to encourage them to be honest when reporting their emotional state. In the cases of students who only interact with PlayPhysics via the web, the ITESM-CCM does not require a participant consent form or formal ethical approval. The only prerequisite that needed to be fulfilled is that lecturers tested, evaluated and approved the GBL environment. They ensured that the physics was modelled accurately and that the instruction provided was accurate, feedback was appropriate and agreed with the learning goals. They also requested changes and improvements that must be implemented before students could interact with PlayPhysics GBL environment, such as setting a message in the GUI for reminding students that they can use the F1 and F2 keys to change between Alpha Centauri's views.

On the other hand, for the version of PlayPhysics GBL environment that includes the GSR Bluetooth sensor or biofeedback device, student consent (Appendix I) needed to be sought, since students must be informed about potential risks before they choose to participate in our experiment. The GSR sensor can cause a mild irritation of the skin, due to the small current applied to measure GSR signal. To diminish this risk, we used the Pre-TENS skin preparation by Tyco (CPR Medical 2012), which reduces skin irritation and enhances electrode adhesion and conductivity. The gel must be applied to each student before connecting him/her to the GSR Bluetooth biofeedback device and, once dried; he/she is connected to it. However, it was expected from the beginning that the group of volunteers willing to participate using the GSR sensor would be fewer in number when compared with the group of volunteers willing to participate interacting with PlayPhysics via the web.

The participants are students enrolled in a related Engineering undergraduate degree or undertaking their last years of high-school, since these students are already considered adults, i.e. they are in the age range of between 18 and 23 years old. Lecturers asked their students to volunteer to participate in the experiment. To encourage them, lecturers offered additional credits in their modules as a reward. In the case of the volunteers who offered to interact using the GSR sensor, they were also offered book vouchers thanks to the support of the e-Learning research group at ITESM-CCM. From the volunteers willing to interact with PlayPhysics via web, lecturers divided students into control and focus groups. The criteria used as a reference to make this division was the overall mark achieved by students at that moment in their modules, the goal being to make both experimental groups as impartial as possible in order to measure the learning gain of students with and without using the GBL environment.

7.1.2 Student participation, tests and samples

Ethical implications and the availability of students per academic semester had also to be considered in order to fulfil the pre-requirements imposed by the institution and the lecturers

and conduct the actual experimentation. At ITESM-CCM, students had three partial exams and one final exam. It is important to observe that students who volunteered were discouraged from participating when they had to study for their exams. Students participated in two phases: (1) Preliminary evaluation and improvements of PlayPhysics GBL environment and (2) Experimentation. In order to receive approval from lecturers in order to conduct experiments with students, lecturers reviewed the appropriateness of PlayPhysics' pre-test, post-test and cut-scenes. Furthermore, lecturers and sixty-six students evaluated and provided comments about PlayPhysics game dialogue in September 2010. Lecturers were also asked to assess PlayPhysics GBL environment from January to March 2011. On this occasion, they invited seventy-nine students to participate in this procedure, who also gave us their comments and suggestions.

Phase	*Testing period*	*Student sample*	*Interacted with*	*Focus*
Preliminary evaluation and improvements of PlayPhysics GBL environment	September 2010	66	PlayPhysics' pre-test and game cut-scene	Evaluation of PlayPhysics' cut-scene
	January-March 2011	79	PlayPhysics online GBL environment	Evaluating PlayPhysics GBL environment.
Experimentation	October-December 2011	118	PlayPhysics online GBL environment	Evaluate *prospective-outcome emotions* network. Define and evaluate *activity and retrospective outcome emotions* networks. Evaluate PlayPhysics GBL environment.
	January-April 2012			
	June-July 2012	8	PlayPhysics on site GBL environment	Define and validate emotional student model with Physiological variables

Table 7.1 Summary of tests conducted with student participants

All the changes given by students and lecturers were assessed (see Appendix J) and the changes that were categorised as having 'High' or 'Medium' priority and that had an impact in functionality and performance were made, otherwise they were ignored. Once changes were made, lecturers allowed us to have access to students in order to conduct tests. For conducting the student interaction with the PlayPhysics GBL environment using the GSR sensor, it was necessary to obtain the approval of the institution and student consent, which involved more time and as a result, these tests were postponed until the final stage of testing. For the

aforementioned reasons, student participation and tests were conducted as summarised in Table 7.1.

The procedure employed to test PlayPhysics online GBL environment in October-December 2011 and January-April 2012 is described in Figure 7.1. Three hundred and ninety-seven students volunteered. Lecturers divided them as objectively as possible into control and focus groups using student performance to date in the module as a reference. Two hundred and two students were allocated to the control group and one hundred and ninety-five to the focus group. Only 118 students in the Focus group had complete data, which entails: solving the pre-test, completing the interaction with PlayPhysics game challenge and solving the post-test. Hence, we needed to select a random sample of 118 students from the two-hundred and two available in the control group in order to conduct a quantitative evaluation of student learning, discussed later. 118 students were in the control and focus group samples respectively. Students in the control group solved the pre-test, learned with a PlayPhysics power point presentation in the form of video and solved the post-test. The random sampling was performed using the filter for resample a data set in the Waikato Environment for Knowledge Analysis (WEKA), a data mining tool created by the University of Waikato (2012).

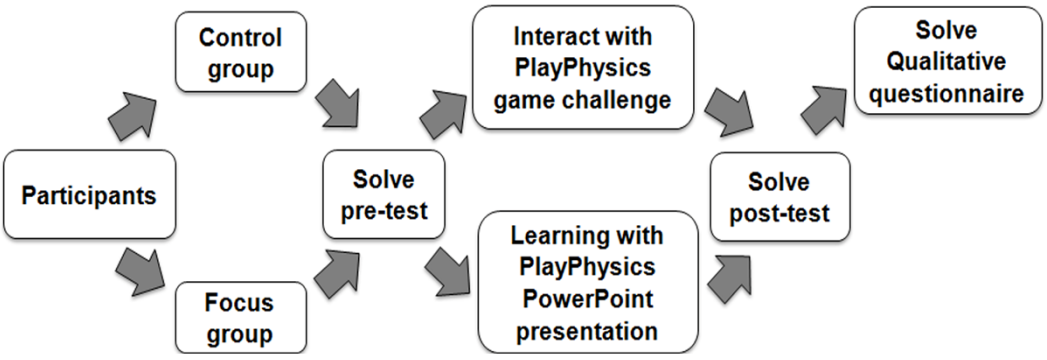


Figure 7.1 Procedure followed to interact with PlayPhysics online GBL environment

The data from students in the focus group was employed to define and evaluate PlayPhysics' emotional student model. Students in both experimental groups answered the pre-test and post-test (Appendix H). However, only students in the focus group interacted with PlayPhysics online GBL environment and were thus eligible to answer the qualitative questionnaire, see Appendix H. Finally 87 out of 118 students answered the qualitative questionnaire. This data is employed to qualitatively evaluate PlayPhysics GBL environment. Students in the control group can only review the PlayPhysics PowerPoint presentation in the form of video, which includes the concepts taught by PlayPhysics GBL environment.

The procedure used to conduct tests on June 2012 is shown in Figure 7.2. Eight students volunteered and signed their consent forms. To test PlayPhysics GBL environment, a room

at ITESM-CCM was acquired and students interacted with PlayPhysics GBL environment on a computer with Internet access, while a member of the e-Learning research team, who assisted us conducting this test, administrated the PC/Server in order to connect the GSR Bluetooth biofeedback device to the student and to monitor the connection. The member of the e-Learning research team also disconnected students at the end of the interaction. Both the student and assistant were located in the same room. After being connected to the GSR sensor, students solved the pre-test, interacted with PlayPhysics GBL environment, solved the post-test and answered the qualitative questionnaire. Finally, students were disconnected from the GSR Bluetooth feedback device. The following section presents the results of the definition and evaluation of PlayPhysics' emotional student model for supporting online educational gaming.

Figure 7.2 Procedure followed to interact with PlayPhysics on-site GBL environment

7.2 Evaluation of PlayPhysics' emotional student model

This section discusses the definition and evaluation of PlayPhysics' emotional student model in PlayPhysics online GBL environment. Performance results corresponding to the prospective outcome, activity and retrospective outcome emotions are presented.

However, before starting to analyse the definition and evaluation of each BBN comprising our emotional model, it is important to define the concept of *time frame t-1* within the context of the dynamic sequence of BBNs implementing our model. The *time frame t-1* can be understood as the instant of time that it is considered the antecedent of the current moment in time that has been recorded in PlayPhysics database every time that the student self-reports his/her emotion or that an event related to student behaviour happens. For example, PlayPhysics game interaction can be conceptualised as a movie tape comprised of specific elements (see Figure 7.3), but its execution is not necessarily in the form of a sequence, since students' decisions/actions determine the order in which the elements are accessed. On the other hand, each achievement emotions network serves a purpose in time, i.e. the *prospective outcome emotions network* is employed for reasoning about emotion in PlayPhysics game dialogues. The *activity emotions network* is employed for reasoning about achievement emotions while students' interact with PlayPhysics game challenge. Finally, the *retrospective outcome emotions network* is used to reasoning about emotion in the instant of time that the outcome of the game challenge is presented to students.

To understand the possible order in which the achievement emotions networks can be used for reasoning about emotion during PlayPhysics execution, we provide the following example. A student starts his/her interaction with PlayPhysics space adventure, the student is presented with a game dialogue introducing the following game challenge and enquiring his/her beliefs or attitudes related to control and value appraisals. At the end of each game dialogue the student self-reports his/her emotion and the *prospective outcome emotions network* can be employed at this moment to reason about emotion (Figure 7.3 (1)). Then the student may proceed to interact with the game challenge, if the student self-reports his/her emotion before completing the game challenge or evaluates the feedback provided by M8 robot, the *activity emotions network* can be used at that instant of time for reasoning about emotion (Figure 7.3 (3)). However, if the latest entry in PlayPhysics database corresponds to a game dialogue, *value t-1* and control *t-1* come from the *prospective outcome emotions network* (Figure 7.3 (1)). Otherwise, if the latest entry corresponds to the ongoing interaction with a game challenge, *value t-1* and *control t-1* come from the *activity outcome emotions network* (Figure 7.3 (2)). However, it is also possible that the latest interaction saved in PlayPhysics database corresponds to the event of notifying students of the outcome, since students can retry the game challenge as many times as desired after receiving their result. Therefore, *value t-1* and *control t-1* can come from the retrospective outcome emotions network (Figure 7.3 (4)). Finally, if students have been presented with their game outcome, they must self-report their emotion towards the received outcome and can then decide whether to proceed with another challenge, thereby starting another game dialogue. In this case *value t-1* and *control t-1* come from the *retrospective outcome emotions network* (Figure 7.3 (5)).

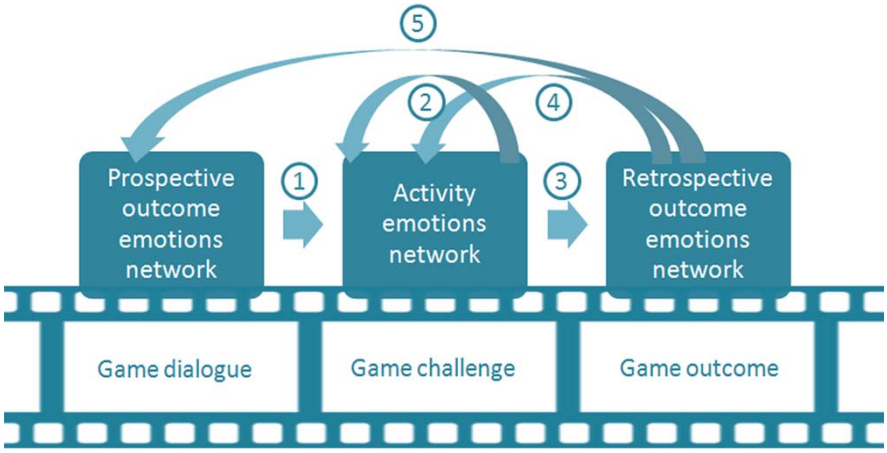


Figure 7.3 Concept of time frame t-1 in emotional student model

It is important to highlight that, in our model, the results of control and value appraisals, which are historic nodes comprising the BBNs, are fed back or passed on (*control t-1* and *value t-1*) instead of the resultant achievement emotion, i.e. *emotion t-1*). The dynamic se-

quence of BBNs resembles the DBN approach, since as mentioned previously, it comprises historic nodes or temporal relationships of the form $Pr(X_t \mid X_{t-1})$. However, it also differs from the DBN approach, since the BBNs also comprise relationships of the form $Pr(X_{t-1} \mid X_t)$.

7.2.1 Prospective outcome emotions

The prospective outcome-emotions are: (1) anticipatory joy, (2) hope, (3) anxiety, (4) anticipatory relief and (5) hopelessness. In addition to these emotions, students were allowed to report feeling no-emotion, i.e. 'Neutral', in order to distinguish the data instances where students felt no emotion, thereby establishing a base line. However, it is also possible that students may report no-emotion when they simply do not wish to disclose their emotional state. The 'Neutral' category is not defined in Control-value theory. Therefore, it represents a limitation for deriving an emotional student model. Pekrun et al. (2007) mention that if control or value is lacking, it cannot be said that the student is feeling a specific emotion. However, it is not always clear when control or value are absent. As a result, if instances where the 'Neutral' emotion was reported by students are included in the final data set used to derive our model, we are interfering with our chances of testing the accuracy of Control-value theory for reasoning about student emotions, since we would be deviating from its definition and would be adding more uncertainty.

To exemplify the above mentioned point, 118 cases were recorded as part of PlayPhysics cut-scene for defining the *prospective outcome* emotions network, which is part of PlayPhysics' emotional student model. In this sample, there are 40 cases of hope, 23 cases of neutral, 21 cases of anticipatory joy, 19 cases of anticipatory relief, 12 cases of anxiety and 3 of hopelessness. Since we cannot assign *control* and *value* for the neutral emotion, we have left these cases empty and we used the 'ReplaceMissingValue' filter in WEKA, which uses the modes and the means to replace the missing values in categorical and numerical variables. Note that *value* is assigned the 'Positive' category and *contro*l is assigned a question mark (?). So in this case, we can assume that *control* appraisal is absent for the *prospective outcome emotions, since neither of the existing categories for control was assigned*.

The three cases of hopelessness were removed, since this number is insufficient to attempt to classify this category. However, if we conduct a comparative analysis using binary and multinomial logistic regression including and having removed the 'Neutral' category and using the Statistical Package for the Social Sciences (SPSS) (IBM 2012), we obtain the results shown in Table 7.2. The potential predictors or independent variables assessed were gender, pre-test, attitude to physics, confidence, source of motivation, perceived difficulty and attitude to effort (see Table 7.5). The continuous variable Pre-test was converted into two categories using the WEKA 'Discretize' filter using equal frequency binning. When value and control were comprised of two categories Binary logistic regression was employed, oth-

erwise Multinomial Logistic Regression (MLR) was applied. For the inclusion of variables, we are employing a criteria from 0.15 to 0.20 suggested by Hosmer and Lemeshow (2000), which ensures the entry of variables with coefficients different to zero, i.e. a probability of 50%. If one of the categories of the independent variables (IVs) achieved this criterion, it was included in the logistic regression equation. Table 7.2 shows the selected predictors.

The logistic equations that we derived are shown in Equations 7.1 to 7.5. Equations 7.1 to 7.3 use a data set that includes the neutral emotion and Equations 7.4 to 7.5 use the same data set, but exclude the neutral emotion. The logistic regression Equations 7.1 and 7.4 show the logit, i.e. natural logarithm of the odds, of having a student with 'negative' value ($p_{negative}$) referenced against the 'positive' value group. In the same manner, Equation 7.2 represents the logit of having a student with 'unknown' value ($p_?$) referenced against 'positive' value group. It is noted that the latter is an additional equation, which must be defined for classifying neutral emotion. It is important to remember that for a unit change in the predictor variable, the logit of outcome (m) relative to the referenced group is expected to change by its respective parameter estimate in logit units, given that the variables in the model are held constant. Equations 7.3 and 7.5 correspond to the logit of having a student with 'high' control (p_{high}) referenced against the 'low' control group. It is noted that when neutral emotion is excluded, the number of cases classified correctly increases for 'high' control and equation 7.5 includes the regressors: attitude to physics and pre-test mark.

Dataset	*Dependent variables (DVs)*	*Predictors (IVs)*	*Significance (p-values)*	*Odds ratios*	*95% Confidence Intervals (C.I.)*	*% cases correctly classified*
Including 'Neutral' emotion (115 cases)	Value 'Negative'	Pre-test '(-inf-87.5])'	0.010	0.194	t0.075 – 0.498	73.9 (85 cases)
		Confidence 'High'	0.108	2.227	0.840 – 5.907	
		Confidence 'Low'	0.234	2.462	0.559 – 10.846	
	Control '?'	Confidence 'High'	0.957	1.032	1.032 – 0.334	51.3 (59 cases)
		Confidence 'Low'	0.087	3.360	0.838 – 13.479	
	Control 'High'	Confidence 'High'	0.021	2.834	1.166 – 6.887	
		Confidence 'Low'	0.869	0.862	0.147 – 5.042	
Excluding 'Neutral' emotion (92 cases)	Value 'Negative'	Pre-test '(-inf-87.5])'	1.864E-4	0.157	0.060 – 0.415	70.7 (65 cases)
	Control 'High'	Attitude to physics 'Negative'	0.892	1.090	0.313 – 3.795	66.3 (61 cases)
		Attitude to physics 'None'	0.042	3.057	1.042 – 8.970	
		Confidence 'High'	0.042	0.346	0.124 – 0.962	
		Confidence 'Low'	0.601	1.656	0.250 – 10.961	
		Pre-test '(-inf-87.5]'	0.075	0.413	0.156 – 1.094	

Table 7.2 Comparison of classification including or excluding the 'Neutral' emotion

$$\ln\left(\frac{p_{negative}}{1-p_{negative}}\right) = 1.484 - 1.642\,pre\_test_{(\text{-inf-87.5}]} + 0.801 confidence_{high} + 0.901 confidence_{low} \quad \text{Eq. 7.1}$$

$$\ln\left(\frac{p_{?}}{1-p_{?}}\right) = -1.030 + 0.031 confidence_{high} + 1.212 confidence_{low} \quad \text{Eq. 7.2}$$

$$\ln\left(\frac{p_{high}}{1-p_{high}}\right) = -0.767 + 1.042 confidence_{high} - 0.149 confidence_{low} \quad \text{Eq. 7.3}$$

$$\ln\left(\frac{p_{negative}}{1-p_{negative}}\right) = 1.658 - 1.849\,pre\_test_{(\text{-inf-87.5}]} \quad \text{Eq. 7.4}$$

$$\ln\left(\frac{p_{high}}{1-p_{high}}\right) = 0.727 + 0.086 attitude\_to\_physics_{negative} + 1.117 attitude\_to\_physics_{none} - 1.062 confidence_{high} + 0.504 confidence_{low} - 0.883\,pre\_test_{(\text{-inf-87.5}]} \quad \text{Eq. 7.5}$$

In order to objectively compare the impact of including the ‘Neutral’ category, in addition to presenting the confusion matrixes in Tables 7.3 to 7.4, we obtained the *sensitivity*, *specificity*, precision and overall accuracy measures of the classifiers, using Equations 7.6 to 7.9 as defined in Han and Kamber (2006). True positives (*t_pos*) are positive tuples that were correctly labelled by the classifier, e.g. value = ’Negative’. True negatives (*t_neg*) are the negative tuples that were labelled by the classifier, e.g. value = ‘Positive’. False positives (*f_pos*) are the negative tuples that were negatively labelled by the classifier, e.g. tuples of the class.value = ‘Positive’ that were classified as value = ‘Negative’. False negatives (f_neg) are the positive tuples that were incorrectly labelled by the classifier. *Sensitivity* is the true positive (*t_pos*) recognition rate. *Specificity* is the true negative (*t_neg*) rate. The *precision* is the percentage of tuples that actually belong to each labelled category. The overall *accuracy* is a function of the *sensitivity* and *specificity*.

$$sensitivity = \frac{t\_pos}{pos} \quad \text{Eq. 7.6}$$

$$specificity = \frac{t\_neg}{neg} \quad \text{Eq. 7.7}$$

$$precision = \frac{t\_pos}{(t\_pos + f\_pos)} \quad \text{Eq. 7.8}$$

$$accuracy = sensitivity\frac{pos}{(pos+neg)} + specificity\frac{neg}{(pos+neg)} \quad \text{Eq. 7.9}$$

As can be observed in Tables 7.3 and 7.4, including instances with the category ‘Neutral’ in the dataset reduces the capability of our classifier to reason accurately about control or value. The sensitivity, the probability of categorising a student with a ‘Negative’ value given that he is focused on the possibility of achieving a failed outcome, is smaller (0.45) for the model including the ‘Neutral’ emotion in comparison to the model excluding it (0.74). So the chances of correctly categorising a student with ‘Negative’ value increases by 0.29 when the ‘Neutral’ category is not included. For the *specificity*, which is the probability of classifying a student with ‘Positive’ value given that he/she is focused on achieving a ‘successful outcome, the probability is also larger when the ‘Neutral’ category is not included. The chances of cor-

rectly classifying a student ‘Positive’ value increases by 0.18. The *precision*, the probability of correctly classifying a randomly selected case corresponding to a student with a ‘Negative’ value, is larger for the classifier excluding the ‘Neutral’ category, i.e. 0.54 instead of 0.52. However, the proportion of true results for value on the population including the ‘Neutral’ category is smaller by 0.03 for the category including the ‘Neutral’ emotion.

Including ‘Neutral’ emotion				
Observed	*Predicted*		*‘Negative’ Sensitivity*	*‘Negative’ Specificity*
	Negative	*Positive*	0.45	0.85
Negative	14	17	*‘Negative’ Precision*	*Model Accuracy*
Positive	13	71	0.52	0.74
Excluding ‘Neutral’ emotion				
Observed	Predicted		‘Negative’ Sensitivity	‘Negative’ Specificity
	Negative	Positive	0.74	0.69
Negative	23	8	‘Negative’ Precision	Model Accuracy
Positive	19	42	0.54	0.71

Table 7.3 Confusion matrixes for *value* including and excluding the ‘Neutral’ emotion

Including ‘Neutral’ emotion					
Observed	*Predicted*			*‘High’ Sensitivity*	*‘High’ Specificity*
	High	*Medium*	*?*	0.63	0.65
High	25	13	2	*‘High’ Precision*	*Model Accuracy*
Medium	19	28	5	0.49	0.51
?	7	10	6		
Excluding ‘Neutral’ emotion					
Observed	*Predicted*			*‘High’ Sensitivity*	*‘High’ Specificity*
	High	*Medium*		0.70	0.63
High	28	12		*‘High’ Precision*	*Model Accuracy*
Medium	19	33		0.60	0.66

Table 7.4 Confusion matrixes for *control* including and excluding the ‘Neutral’ emotion

On the other hand, the *sensitivity*, which is the probability of classifying the control of a student as ‘High’ given that he/she has a positive perspective about his/her skills is 0.7 for the classifier excluding the ‘Neutral’ emotion, while for the classifier including the ‘Neutral’

emotion this probability is decreased by 0.07. The *specificity*, the probability of correctly classifying the control of a student as 'Medium' when he/she believes that his/her skills match this level, decreases by 0.02 when the 'Neutral' category is not included. The *precision* of correctly classifying a randomly selected case corresponding to a student with 'High' control is increased by 0.11 as the overall *accuracy* of the classifier, which in this case increases by 0.15.

		Predictors including the 'Neutral' category	*Predictors excluding the 'Neutral' category*
Predictor	*Category*	*N° cases*	*N° cases*
Gender	F	41	33
	M	74	59
Pre-test	(-inf-87.5]	52	42
	(87.5-inf)	63	50
Attitude to physics	Negative	28	20
	None	43	37
	Positive	44	35
Confidence	High	51	44
	Low	13	7
	Medium	51	41
Source of motivation	External/Both	48	37
	Inner	67	55
Perceived difficulty	Low/High	51	35
	Medium	64	57
Attitude to effort	None/Negative	22	16
	Positive	93	76

Table 7.5 Potential predictors of *prospective outcome emotions*

As a result, we employed 92 cases to define, train and evaluate the *prospective outcome emotions* network. The structure of the *prospective outcome emotions* network was created by applying the Necessary-Path Condition algorithm with a level of significance of 0.05 in Hugin Lite (Hugin Expert A/S 2011) in combination with knowledge of the domain acquired through employing Binary Logistics regression in SPSS. The EM learning algorithm in Hugin Lite was employed to perform parameter learning (see Appendix K). The prospective outcome emotions network derived after employing all the available data is shown in Figure 7.4. Ten-fold cross-validation is employed to determine how this model will perform over future data. Ten stratified folds were extracted from the original data set of 92 instances. To create the ten data folds, a small program was implemented using PHP. Two programs in Java, BayesianNetNov11Jun12ThresholdReducFinal and TestBeforeNov11Jun12ThresholdReduFinal, which use the Hugin Lite API (HAPI), were created to propagate the evidence and ob-

tain the predictions of the network. For the prospective outcome emotions network, ninety predictions and nine different structures were obtained in total by applying the Necessary-Path Condition method and resolving the uncertain relations using the knowledge obtained from applying the Binary Logistic regression and the Pearson correlations (see Appendix K).

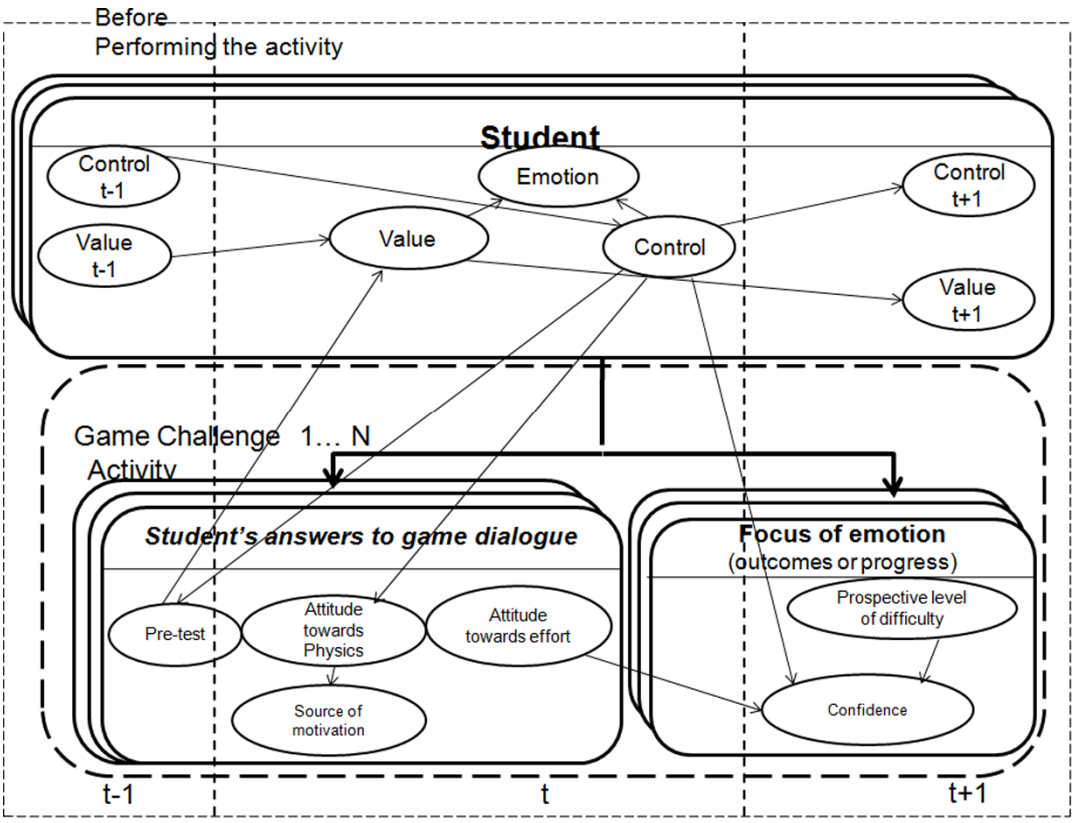


Figure 7.4 PlayPhysics *prospective outcome* emotions network

The performance of the *prospective outcome emotions* network over all the training data-set is presented in Tables 7.6 to 7.8. This network is capable of predicting the student *value* with a probability of 0.80 (see Table 7.7), i.e. whether the student is focused on the likelihood of achieving a successful ('Positive') or failed ('Negative') outcome. Also, this same network is capable of predicting the student *control* with a probability of 0.6667 (see Table 7.8), i.e. whether the student considers that his/her skills correspond to 'High' or 'Medium' control over the specific game challenge in order to achieve a successful outcome. However, the network is capable of predicting *prospective outcome* achievement emotions with a probability of 0.5667 (see Table 7.6), which accurately classifies approximately 57% of the cases.

Prospective outcome achievement emotions have a Pearson correlation of 0.224 with *value* that is significant at the 0.01 level (2-tailed), i.e. r(92) = 0.224. The square of the coefficient of determination (r^2) is employed to analyse the effect size, in this case r^2 = 0.050. Following Cohen's (1988) criteria for classification of association strength, since r^2 is less than 0.09 and larger than 0.01, the association is *small* and between 1% and 8% of the variance

is shared. On the other hand, *prospective outcome achievement emotions* have a Pearson correlation of 0.928 that is significant at the 0.05 level (2-tailed), i.e. r(92) = 0.928 and r^2 = 0.861, since r^2 is larger than 0.25 the effect is large and we can state that at least 25% of the variance is shared. As a result, it can be said that the likelihood of accurately classifying *prospective outcome achievement emotions increases* when *control* is classified correctly.

The *prospective outcome emotions* network will accurately classify emotions that depend on the accurate classification of 'Medium' control, such as Hope and Anxiety which were identified with 67.5% and 80% accuracy respectively. Also, the past-outcome, i.e. pre-test mark, is a reasonable regressor of student value, i.e. the likelihood of focusing on the possibility of a successful or failed outcome during the interaction with PlayPhysics' game challenge. 'Medium' control is also reasonably predicted by student confidence, attitude towards physics and past experience, e.g. pre-test. However, it is necessary to identify another regressor for 'High' control in order to improve the accuracy of the outcome-prospective emotions network. Even though the random variable gender was not selected for the final prospective outcome emotions network, it was noted that during cross-validation that it was sometimes associated with student confidence and student attitudes towards physics. The latter agrees with previous findings in Muñoz et al. (2012), where prospective outcome emotions were classified in two categories, *positive* and *negative* emotions. In that case, it was found that, from a population of 40 males and 26 females, the association between *gender* and the attitude towards physics was negative and small, i.e. r(66) = -0.284 and r^2 = 0.081. 57.7% from the male sample reported having a neutral or negative attitude towards physics in comparison to 84.6% from the female sample, who reported a neutral or negative attitude towards physics. In this case, there are 33 females and 59 males and the association between gender and student confidence is negative and medium, i.e. r(92) = -0.332 and r2 = 0.110. Hence, approximately between 1% and 8% of the variance is shared between these two variables. From this sample 61% from the male sample reported having 'High' confidence, 33.9% reported having a 'Medium' level of confidence' and 5.1% reported 'Low' confidence on achieving a successful gaming and learning outcome, whilst 24.2% from the females sample reported having 'High' confidence, 63.3% reported having a 'Medium' level of confidence and 12.1% reported having 'Low' confidence on a achieving a successful outcome. So it can be said that female students require more reassurance and encouragement than male students during instruction.

If we look at the *sensitivity* in Table 7.6, we can observe that the *prospective outcome emotions* network has varying ability to accurately predict certain classes (true positives). For example, *Anxiety* is predicted with a probability of 0.8, which means that 80% of the cases corresponding to this class are classified correctly, whilst *Anticipatory Joy* is predicted with a probability 0.3, which corresponds to 30% of the cases in this class. In the same manner the

network has varying values for the *precision* and the *specificity* of each category. The former is related to the probability of classifying an instance as belonging to the Anxiety category given that the student is feeling actually anxious, while the latter measures the true negative rate, i.e. the probability of classifying the instances as belonging to Hope, Anticipatory Joy and Anticipatory Relief categories, given that the student does not actually feel Anxiety.

Prospective outcome emotions								
Observed	Predicted							
	Anticipatory Joy	Anticipatory Relief	Anxie-ty	Hope	Specificity	Sensitivity	Preci-sion	Network accuracy
Anticipatory Joy	6	3	1	10	0.857	0.300	0.375	0.567
Anticipatory Relief	2	10	2	6	0.914	0.500	0.625	
Anxiety	0	1	8	1	0.925	0.800	0.571	
Hope	8	2	3	27	0.660	0.675	0.613	

Table 7.6 Confusion matrix corresponding to *prospective outcome emotions*

Value of *prospective outcome* emotions						
Observed	Predicted					
	Negative	Positive	Specificity	Sensitivity	Precision	Network accuracy
Negative	21	9	0.850	0.700	0.700	0.800
Positive	9	51	0.700	0.850	0.850	

Table 7.7 Confusion matrix corresponding to *value*

Control of *prospective outcome* emotions						
Observed	Predicted					
	High	Medium	Specificity	Sensitivity	Precision	Network accuracy
High	21	19	0.780	0.520	0.660	0.670
Medium	11	39	0.520	0.780	0.670	

Table 7.8 Confusion matrix corresponding to *control*

For the aforementioned reasons, reporting the percentages of agreement between the student self-report and PlayPhysics *prospective outcome emotions* network of 56.67% (probability of 0.5667), 66.67% (probability of 0.6667) and 80% (probability of 80%) for *emotion*, *control* and *value* respectively can be misleading in this situation, since there are varied frequencies for each column and row in the confusion matrices or contingency tables. These differences show that there was a tendency for choosing certain diagnostic categories more frequently than others. Therefore, to determine how closely the student self-reports agreed

with the predictions for *emotion*, *control* and *value* of PlayPhysics' *prospective outcome emotions* network*, Cohen's Kappa, a form of intra-class correlation coefficient* is applied (Landis and Koch 1977, Cohen 1960). The Kappa (κ) coefficient is employed as a measure of agreement that adjusts the observed proportional agreement by considering the amount of agreement that would be expected by chance. Cohen's Kappa is presented in Equation 7.10 (Cohen 1960), where p is the proportion of instances where there is an agreement and p_e is the proportion of instances that would be expected to agree by chance. Cohen's Kappa is calculated in this dissertation in SPSS as shown in Equation 7.11 (Kinnear and Gray 2010). If we have ordered our observed and predicted values in a contingency table, *O* and *E* are the observed and predicted or expected frequencies respectively for the cells in the main diagonal and N is the total number of cases or instances. Kappa can take values in a range [-1, 1], but only values in a range [0, 1] are meaningful. According to Landis and Koch (1977) the degree of agreement of Kappa for hypothesis testing is judged as summarised in Table 7.9.

$$\kappa = \frac{p - p_e}{1 - p_e} \qquad \text{Eq. 7.10}$$

$$\kappa = \frac{\sum_{diagonal} O - \sum_{diagonal} E}{N - \sum_{diagonal} E} \qquad \text{Eq. 7.11}$$

Range of Kappa	*Judgement*
$if(\kappa < 0)$	"There is poor or no agreement"
$if(\kappa > 0)\,\&\,(\kappa \le 0.2)$	"There is a slight agreement"
$if(\kappa > 0.2)\,\&\,(\kappa \le 0.4)$	"There is a fair agreement"
$if(\kappa > 0.4)\,\&\,(\kappa \le 0.6)$	"There is a moderate agreement"
$if(\kappa > 0.6)\,\&\,(\kappa \le 0.8)$	"There is a substantial agreement"
$if(\kappa > 0.8)\,\&\,(\kappa \le 1)$	"There is almost a perfect agreement"

Table 7.9 Judgment criteria for Cohen's Kappa

Accordingly, our hypotheses are stated as follows:

H_0 - The degree of agreement between the student self-report and *PlayPhysics' prospective outcome emotions* network classification is random (κ=0)

H_A - There is a reasonable degree of agreement between the student self-report and PlayPhysics' *prospective outcome emotions* network classification

The results of 90 testing cases are presented in Table 7.10. Cohen's Kappa shows a 'fair agreement' for *emotion* and *control*, i.e. κ=0.369; p<0.01 and κ=0.311; p<0.01 respectively. For *value*, Cohen's Kappa shows a 'moderate agreement', i.e. κ=0.550; p<0.01. Accordingly,

we reject the null hypothesis and it can be said that the result was not random and that the *prospective outcome emotions* network is giving a fair to moderate accuracy of classification. However, its accuracy shows opportunity for enhancement in further work, since diagnostic systems that are considered *highly* reliable give at least κ=0.75 (Bouckaert et al. 2012, Kinnear and Gray 2010).

Cohen's Kappa for the prospective outcome achievement emotions network		
Dependent variable	*K*	*Significance*
Emotion	0.369	7.732E-9
Value	0.550	1.811E-7
Control	0.311	0.003

Table 7.10 Cohen's Kappa for the prospective outcome emotions network

7.2.2 Activity emotions

From 118 students, there were 946 cases or instances corresponding to their interaction with PlayPhysics game challenge. However, after filtering the cases corresponding to the neutral emotion, we obtained 708 instances, where 136 instances were reported as Anger, 122 instances were reported as Boredom, 262 instances as Enjoyment and 188 as Frustration. Before filtering the cases corresponding to Neutral emotion, we employed the unsupervised attribute filter 'ReplaceMissingValues' of WEKA to assign the corresponding values of *control* and *value* to this category, which replaced the missing values using the modes and means of the attributes in the dataset. In this case, the WEKA filter assigns the following values for the Neutral category: 'High' for control and '?' for value. The latter may indicate the possibility of an additional category for the dependent variable *value* that should be considered for representing *activity emotions*. As mentioned earlier, Neutral is not employed, since we want to follow Control-value theory as closely as possible. For the creation of the *activity achievement emotions* network, the categories 'High' and 'Low' for the dependent variable *control* and 'Positive', 'None' and 'Negative' for the dependent variable *value* are considered.

We are employing the free version for proof of concept of Hugin Lite to derive the *activity emotions* network. As a result, we are limited to 50 states and 500 cases or instances and we must perform stratified random sampling using the WEKA supervised filter 'Resample' over our data set of 708 cases in order to obtain 499 cases to create the *activity achievement* emotions network in Hugin Lite and perform cross-validation. However, in order to analyse the most suitable manner to discretise the continuous independent variables and to decide the variables that should be included, we performed Binary Logistic Regression for *control* and Multinomial Logistic Regression for *value using* 708 cases. The conversion between continuous and categorical variables was performed with the unsupervised filter 'Discretize' in WEKA. The independent variables employed are *outcome, times asked help, attempts*

alone, estimated independence, overall attempts, average quality tutoring feedback, focus coarse value, interval of interaction and time to achieve learning goals. These variables were discretised into two and three categories using equal frequency binning, since *control* and *value* are originally divided into two and three categories respectively and we have the constraint for representing our model using 50 states in Hugin Lite.

When the confusion matrixes were compared, it was observed that the division into two categories including the independent variables *control t-1* and *value t-1* was the most suitable for classifying *control,* achieving approximately 70.3% accuracy in the cases of 'High' and 'Low' *control* with 0.774 and 0.613 sensitivity respectively. However, the division into three categories was the most suitable for classifying *value* achieving 61.7% accuracy in the cases and '*Negative', 'None'* and *'Positive' value* with 0.769, 0.016 and 0.71 sensitivity (see Appendix K). However, we decided to employ the dataset divided into two categories as shown in Table 7.11, since after reviewing the Pearson correlations of the dataset with 499 cases (Appendix K), it was noted that activity emotions correlate 0.409 at 0.05 level (2-tailed) with *control,* i.e. r(499) = 0.409 and r^2 = 0.167, and correlate 0.117 at 0.05 level (2-tailed) with *value*, i.e. r(499) = 0.117 and r^2 = 0.014. As a result, the association that *control* has with activity emotions is *medium*, while the association that *value* has is *small*. If we want to increase our chances of correctly predicting *activity* emotions, we have to ensure that *control* has the best accuracy possible. *Control t-1* and *value t-1* may come from the *prospective outcome emotions*, *activity emotions* and *retrospective outcome emotions* networks. Therefore, *control t-1* can comprise states such as 'High' and 'Self' simultaneously. Timestamps of the student interactions were employed as a reference to assemble the dataset to derive this network.

Tables 7.12 to 7.13 shows the predictors selected after applying Binary and Multinomial Logistic Regression over 499 cases using the 'Conditional Forward' or 'Stepwise' procedure for *control* and *value* respectively. Equations 7.12 to 7.14 are the logistic regression equations derived by applying the afore-mentioned procedures. Equation 7.12 represents the logit of having a student with 'high' control (p_{high}) referenced against the 'low' control group. Equations 7.13 to 7.14 represent he logit of having a student with 'negative' value ($p_{negative}$) and having a student with 'none' value (p_{none}) respectively, in both cases referenced against the 'positive' value group. While comparing the regression equations defined for the prospective outcome emotions and the activity emotions, it was noted that an additional equation is necessary to classify activity emotions. This information in conjunction with the Pearson correlations (see Appendix K) is employed to solve uncertain associations when applying the Necessary-Path Condition algorithm to create the structure of the *activity emotions* network. EM learning is again employed to perform parameter learning (see Appendix K). The activity

emotions network derived from 499 instances is shown in Figure 7.5. There are temporal relations corresponding to the past time slice and emotion, e.g. control t-1 and value t-1.

Predictor	*Category*	*N° cases*	*Predictor*	*Category*	*N° cases*
Outcome	(-inf-9.285]	6	Focus coarse value	(-inf-1.765]	76
	(9.285-inf)	493		(1.765-inf)	423
Type of outcome	Black out	1	Interval of interaction	(-inf-559s]	265
	In progress	492		(559-inf)	234
	Positive acceleration	6	Focus coarse value	(-inf-2.805]	232
Times asked help	(-inf-0.5]	255		(2.805-inf)	267
	(0.5-inf)	244	Time to achieve learning goals	(-inf-1242.301724s]	464
Attempts alone	(-inf-1.5]	284		(1242.301724s-inf)'	35
	(1.5-inf)	215	Control t-1	High	201
Estimated independence	(-inf-0.5]'	260		Irrelevant	32
	(0.5-inf)	239		Low	157
Overall attempts	(-inf-1.5]	227		Medium	19
	(1.5-inf)	272		Other	52
Value t-1	Negative	287		Self	38
	None	52			
	Positive	160			

Table 7.11 Potential predictors of *activity emotions*

Dependent variables (DVs)	*Predictors (IVs)*	*Significance (p-values)*	*Odds ratios*	*95% Confidence Intervals (C.I.)*	*% cases correctly classified*
Control 'High'	Outcome '(-inf-9.285]'	0.094	6.737	0.721 - 62.972	70.7 (353 cases)
	Attempts alone '(-inf-1.5]'	0.011	1.701	1.127 – 2.567	
	Average tutoring feedback '(-inf-1.765]'	0.004	2.283	1.296 – 4.022	
	Value t-1 'Negative'	0.016	1.945	1.134 – 3.338	
	Value t-1 'None'	0.566	1.303	0.528 – 3.216	
	Control t-1 'High'	0.194	0.601	0.278 – 1.296	
	Control t-1 'Irrelevant'	0.138	2.117	0.786 – 5.703	
	Control t-1 'Low'	0.010	2.878	1.290 – 6.423	
	Control t-1 'Medium'	0.646	0.742	0.208 – 2.646	
	Control t-1 'Other'	0.148	0.512	0.207 – 1.268	

Table 7.12 Predictors selected for *control* using Binary Logistic Regression

$$\ln\left(\frac{p_{high}}{1-p_{high}}\right) = -1.194 + 1.908 outcome_{(-\inf -9.285]} + 0.531 attempts\_alone_{(-\inf -1.5]} - 0.510 control\_t\_1_{high}$$
$$+ 0.826 average\_tutoring\_feedback_{(-\inf -1.765]} + 0.665 value\_t\_1_{negative} + 0.265 value\_t\_1_{none}$$
$$- 0.298 control\_t\_1_{medium} + 0.750 control\_t\_1_{irrelevant} - 0.670 control\_t\_1_{other} + 1.057 control\_t\_1_{low}$$

Eq. 7.12

Dependent variables (DVs)	*Predictors (IVs)*	*Significance (p-values)*	*Odds ratios*	*95% Confidence Intervals (C.I.)*	*% cases correctly classified*
Value 'Negative'	Outcome '(-inf-9.285]'	0.735	0.605	0.033 - 11.126	60.3 (301 cases)
	Times asked help '(-inf-0.5]'	0.006	0.518	0.324 - 0.829	
	Value t-1 'Negative'	2.11E-7	4.793	2.712 - 8.848	
	Value t-1 'None'	0.891	0.912	0.246 - 3.377	
	Control t-1 'High'	0.217	1.706	0.731 - 3.985	
	Control t-1 'Irrelevant'	0.978	0.984	0.308 - 3.145	
	Control t-1 'Low'	4.34E-5	11.384	3.547 - 36.531	
	Control t-1'Medium'	0.903	1.102	0.232 – 5.242	
	Control t-1 'Other'	0.863	0.920	0.358 – 2.365	
Value 'None'	Outcome '(-inf-9.285]'	0.097	7.339	0.696 – 77.354	
	Times asked help '(-inf-0.5]'	0.435	0.794	0.446 – 1.415	
	Value t-1 'Negative'	0.001	3.654	1.720 – 7.766	
	Value t-1 'None'	0.739	1.280	0.300 – 5.456	
	Control t-1 'High'	0.188	2.264	0.671 – 7.644	
	Control t-1 'Irrelevant'	0.033	4.572	1.134 – 18.425	
	Control t-1 'Low'	2.15E-4	15.528	3.634 – 66.353	
	Control t-1 'Medium'	0.171	3.297	0.596 – 18.230	
	Control t-1 'Other'	0.569	0.642	0.140 – 2.948	

Table 7.13 Predictors selected for *value* using Multinomial Logistic Regression

$$\ln\left(p_{negative}/1-p_{negative}\right) = -1.013 - 0.502outcome_{(\text{-inf-9.285]}} - 0.658times\_asked\_help_{(\text{-inf-0.5]}} + 1.567value\_t\_1_{negative} - 0.092value\_t\_1_{none} + 0.534control\_t\_1_{high} - 0.016control\_t\_1_{irrelevant} + 2.432control\_t\_1_{low} + 0.097control\_t\_1_{medium} - 0.083control\_t\_1_{other} \quad \text{Eq. 7.13}$$

$$\ln\left(p_{none}/1-p_{none}\right) = -2.347 + 1.993outcome_{(\text{-inf-9.285]}} - 0.230times\_asked\_help_{(\text{-inf-0.5]}} + 0.247value\_t\_1_{none} + 0.817control\_t\_1_{high} + 1.520control\_t\_1_{irrelevant} + 2.743control\_t\_1_{low} + 1.193control\_t\_1_{medium} - 0.444control\_t\_1_{other} + 1.296value\_t\_1_{negative} \quad \text{Eq. 7.14}$$

The activity emotions network was validated using stratified 10-fold cross-validation. As a result 500 cases comprised the testing data set. Tables 7.14 to 7.16 show results corresponding to these tests presented in the form of confusion matrices. It can be observed that the activity emotions network predicts the student *control* while learning and playing with PlayPhysics game challenge with a probability of 0.72 (see Table 7.16). The student *value*,

whether the student has a 'Positive', 'None' or 'Negative' perception of the game activity, is predicted with a probability of 0.63 (see Table 7.15).

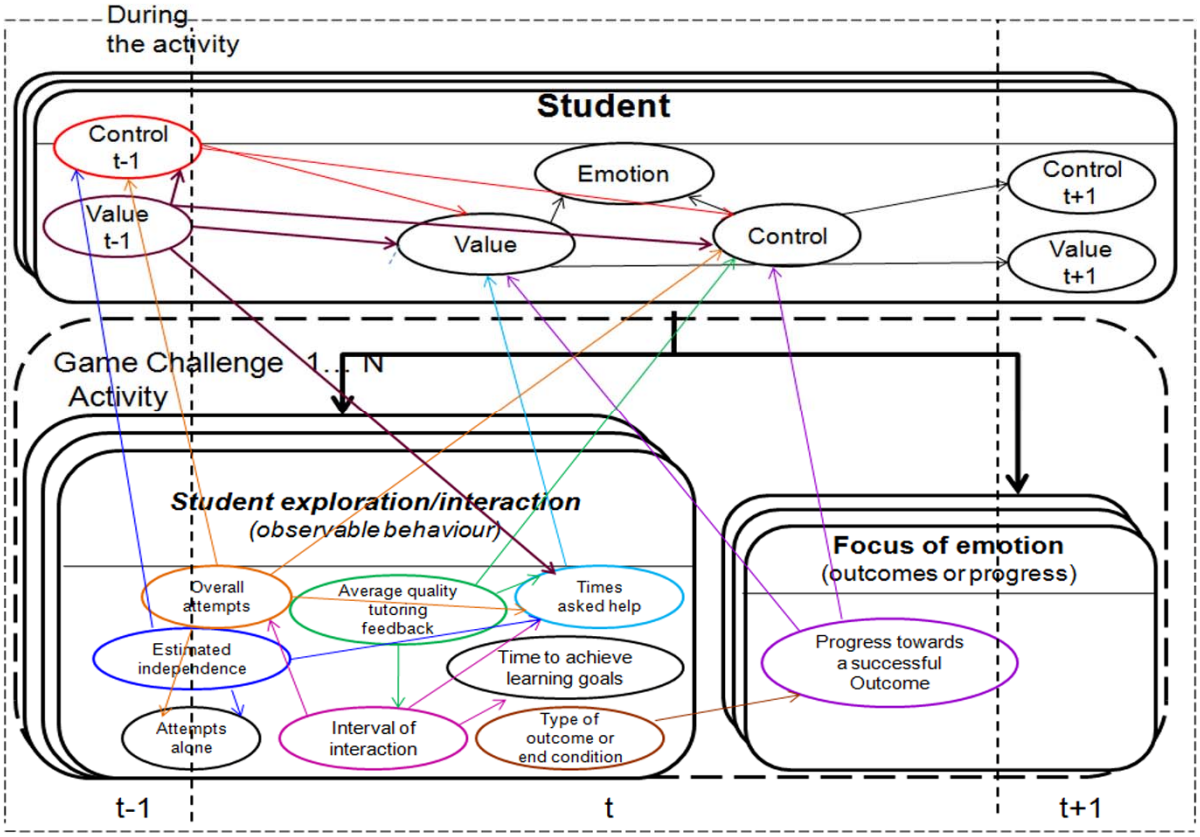


Figure 7.5 *Activity achievement emotions* network

However, *activity emotions* are predicted with a probability of 0.53 (see Table 7.14), where the achievement emotions, enjoyment and frustration, report a higher number of cases accurately classified, with sensitivity probabilities of 0.678 and 0.600 respectively. The cases also have higher probabilities of precision (0.595 and 0.523). *Boredom* has the least classification accuracy. This is due to not being able to recognise *value* 'None' with the available independent variables. *Control* and *value* are shown in the network structure as simultaneously influenced by *control t-1* and *value t-1*. The association of *value t-1* with the current *value* is medium and significant at 0.05 level (2-tailed), i.e. $r(499) = 0.387$ and $r^2 = 0.149$. Whilst the association between *value t-1* and *control t-1* with *control* is small, $r(499) = 0.121$; $r^2 = 0.015$; $p<0.01$ and $r(499) = 0.253$; $r^2 = 0.064$, $p<0.05$ respectively. *Value* is also shown in the network as influenced by the progress towards the game outcome and the number of times that the student asked help from M8 robot. *Control* is shown in this network as determined by students' overall attempts, the progress towards the outcome and student perception of the quality of hints, i.e. average quality of the tutoring feedback provided by M8 robot.

Also, it is noted that the student interval of interaction with PlayPhysics depends on the student overall attempts, the time that students take to achieve the learning and game goals

and the number of times students enquired help from M8-robot. Even though the random variable *focus coarse value,* related to the *mouse focus,* was not selected for the final network, it was sometimes associated with *control* during cross-validation. In addition, it was noted that *control* 'High' and *value* 'Negative' have the highest quantity of cases correctly classified, with probabilities of sensitivity of 0.82 and 0.757. However, the emotion enjoyment has a higher probability of sensitivity, even though its classification depends on 'Positive' *value*. This may be is due to the manner in which *activity* emotions interact. Activity emotions include two emotions – anger and frustration - which depend on the correct classification of 'Negative' *value* and also include –anger and enjoyment – which depend on the correct classification of 'High' control. As a result, it may be that as we increase the likelihood of recognising enjoyment and frustration, we are reducing the likelihood of correctly classifying anger.

Activity emotions								
Observed	Predicted							
	Anger	Boredom	Enjoyment	Frustration	Specificity	Sensitivity	Precision	Network accuracy
Anger	48	2	27	23	0.848	0.480	0.440	0.532
Boredom	15	18	31	26	0.954	0.200	0.486	
Enjoyment	30	6	122	22	0.740	0.678	0.595	
Frustration	16	11	25	78	0.808	0.600	0.523	

Table 7.14 Confusion matrix corresponding to *activity emotions*

Value of *activity* emotions							
Observed	Predicted						
	Negative	None	Positive	Specificity	Sensitivity	Precision	Network accuracy
Negative	174	10	46	0.641	0.757	0.642	0.630
None	47	17	26	0.961	0.189	0.515	
Positive	50	6	124	0.775	0.689	0.633	

Table 7.15 Confusion matrix corresponding to *value*

Control of *activity* emotions						
Observed	Predicted					
	High	Low	Specificity	Sensitivity	Precision	Network accuracy
High	230	50	0.600	0.800	0.720	0.720
Low	88	132	0.820	0.600	0.730	

Table 7.16 Confusion matrix corresponding to *control*

In this case, as in the evaluation of the prospective outcome emotions network, we employ Cohen's Kappa to assess how closely the student self-reports agreed with the predictions for

activity emotions (Landis and Koch 1977, Cohen 1960). Accordingly, we also state our hypotheses as:

H_0- The degree of agreement between the student self-report and *PlayPhysics' activity emotions* network classification is random (κ=0)

H_A- There is a reasonable degree of agreement between the student self-report and PlayPhysics' *activity emotions* network classification

Table 7.17 shows Cohen's Kappa for *control*, *value* and emotion for 500 test instances. Along with Landis and Koch (1977)'s criteria of agreement for Kappa, Cohen's Kappa shows a 'fair agreement' for *emotion* and *value*, i.e. κ=0.348; p<0.01 and κ=0.381; p<0.01 respectively. In this case, Cohen's Kappa shows a 'moderate agreement' with *control*, i.e. κ=0.429; p<0.01. As a result, we also reject the null hypothesis and it also can be said that results obtained were not random, since the *activity emotions network* has a fair to moderate accuracy of classification.

Cohen's Kappa for the activity achievement emotions network		
Dependent variable	*K*	*Significance*
Emotion	0.348	5.108E-39
Value	0.381	2.726E-29
Control	0.429	2.433E-22

Table 7.17 Cohen's Kappa for the activity emotions network

7.2.3 Retrospective outcome emotions

In total, there were collected 516 registries corresponding to 118 students who interacted with PlayPhysics' game-base learning environment for determining *the retrospective outcome emotions* network. However, after filtering the cases corresponding to student self-reports of neutral emotion, we obtained 259 cases. However, when we employed the WEKA unsupervised attribute filter 'ReplaceMissingValues', which uses the modes and the means of each independent variable to replace all missing numerical and categorical data, it substituted the 'Neutral' emotion by a question mark symbol '?' in *value* and 'Other' in *control.* So it can be inferred that an additional category is required for classifying the neutral emotional state using the *retrospective outcome emotions* network. From the final 259 cases, 40.2% correspond to anger, 5.4% from the cases correspond to gratitude, 8.9% from the cases correspond to joy, 12% from cases correspond to pride, 17.8% from cases correspond to sadness and 41% from the cases correspond to shame. Since we are employing the Hugin Lite free version for proof of concept, we also have to consider the constraints of 500 cases and 50 states to define our model.

Predictor	*Category*	*N° cases*	*Predictor*	*Category*	*N° cases*
Outcome	(-inf-55]	136	Focus coarse value	(-inf-2.55]	130
	(55-inf)	123		(2.55-inf)	129
Type of outcome	Black out	49	Time to achieve learning goals	(-inf-1981.664s]	68
	DistanceTooFar	15		(1981.664s-inf)	191
	DistanceTooShort	41	Publishing outcome	Published	35
	Positive acceleration	68		Unpublished	224
	Re-started	1	Interval of interac-tion	(-inf-773s]	130
	Success	66		(773s-inf)	129
	TimeOut	19	Control t-1	High	83
Times asked help	(-inf-0.5]	112		Irrelevant	28
	(0.5-inf)	147		Low	59
Attempts alone	(-inf-2.5]	149		Medium	11
	(2.5-inf)	110		Other	51
Estimated inde-pendence	(-inf-1.5]	151		Self	27
	(1.5-inf)	108	Value t-1	Negative	142
Overall attempts	(-inf-2.5]	119		None	29
	(2.5-inf)	140		Positive	88
Average quality tutoring feedback	(-inf-1.53]	50			
	(1.53-1.595]	180			
	(1.595-inf)	29			

Table 7.18 Potential predictors of *retrospective outcome emotions*

For the retrospective outcome emotions, we have three states –'Other', 'Self' and 'Irrelevant' for *control* and two states – 'Positive' and 'Negative' for *value*. Hence, to decide whether it is more suitable to discretise the continuous independent variables into two or three categories, we performed an analysis using Binary and Multinomial Logistic Regression (see Appendix K). During this analysis, it was observed that for classifying the dependent variable *control*, the classifier dividing the independent variables or regressors into three categories and including *control t-1* is the most suitable logistic regression model. However, if we divided all the continuous variables in our model into three categories, the variables would not fit in Hugin lite, since there would be approximately 56 states. We identified that the independent variable *average quality of tutoring feedback* divided into three categories is responsible for the improvement. Hence, we decided to divide this variable into three categories and the other variables into two categories. The final dataset employed is shown in Table 7.18. The rationale behind deciding to divide continuous variables into three and two categories lies in knowing that retrospective outcome emotions have a Pearson correlation of 0.163 at 0.01 level (2-tailed), r(259) = 0.163 and r^2 = 0.027, which means that the association between *control* and the *retrospective outcome* emotions is small, i.e. they only share be-

tween 1% and 8% of the variance. The relation between *retrospective outcome* emotions and value is not significant. However, *retrospective outcome emotions* have a small significant relation at 0.05 level (2-tailed) with *value t-1*, r(259) = 0.145 and r^2 = 0.021. Hence, we focus on enhancing the classification of *control.*

Tables 7.19 to 7.20 show the random variables selected after applying Binary and Multinomial Logistic Regression (BLR/MLR) for value and control respectively using 259 cases and the 'Conditional Forward' or 'Stepwise' procedure. Equations 7.15 to 7.17 were derived using the previously mentioned approaches. Equation 7.15 shows the logit of having a student with 'negative' value ($p_{negative}$) referenced against the 'positive' value group, while Equations 7.16 to 7.17 present the logit of having a student with 'irrelevant' control ($p_{irrelevant}$) and having a student with 'other' control (p_{other}) respectively, both referenced against the 'self' control group. It is noted that to classify retrospective outcome emotions, it is necessary to define two equations for reasoning about student control.

This information in combination with the information of the Pearson correlations (see Appendix K) is employed to apply the Necessary-Path Condition algorithm for structured learning and solving the uncertain relations. Also on this occasion, we employed the EM learning algorithm to perform parameter learning. The resultant retrospective outcome emotions network, obtained employing 259 cases, is shown in Figure 7.6. The EM learning algorithm was also employed for learning the Conditional Probability Tables (CPTs), see Appendix K. From this network, it is noted that *value* is influenced by *value t-1*, *control t-1*, whether or not students decided to publish their result and the type of outcome or end condition. The latter indicates whether the outcome was successful, the type of misconception or whether the student decided to Re-start the game challenge or quit. When the Pearson correlation coefficients were reviewed for *control t-1*, *value t-1* and the *type of outcome,* they were shown to have a small size association (r(259) = 0.221;p<0.01, r(259) = -0.133; p<0.05 and r(259) =0.223; p<0.01 respectively) sharing between 1% or 8% of the variance with *value.* However, the random variable *publishing outcome* has a large association with *value* (r(259) =-0.508; p<0.01). On the other hand, control is related to *control t-1*, which can come from the *activity emotions* network or from the retrospective outcome emotions network. Also, control was related to the student willingness to publish the outcome, which demonstrated to be a negative and small size association (r(259) = -0.177; p<0.01). The student *control* was also linked to the average quality of tutoring feedback. For the evaluation of the *retrospective outcome emotions* network, we applied 10-fold cross-validation. Results corresponding to 260 test instances are presented in Tables 7.21 to 7.23.

Dependent variables (DVs)	*Predictors (IVs)*	*Significance (p-values)*	*Odds ratios*	*95% Confidence Intervals (C.I.)*	*% cases correctly classified*
Value 'Negative'	Type of outcome 'BlaOut'	0.108	6.247	0.668 - 58.418	83.8 (217 cases)
	Type of outcome 'DistanceTooFar'	0.941	0.892	0.044 – 17.999	
	Type of outcome 'DistanceTooShort'	0.066	8.377	0.869 – 80.781	
	Type of outcome 'PositiveAcceleration'	0.427	2.520	0.258 – 24.616	
	Type of outcome 'ReStarted'	1.000	7.44E-8	0.000 -	
	Type of outcome 'Success'	0.025	13.357	1.395 – 127.925	
	Publishing outcome 'Published'	1.52E-4	11.305	3.223 – 39.658	
	Control t-1 'High'	0.633	0.713	0.177 – 2.865	
	Control t-1 'Irrelevant'	0.048	4.806	1.011 – 22.845	
	Control t-1 'Low'	0.108	3.956	0.738 – 21.209	
	Control t-1 'Medium'	0.918	0.894	0.105 – 7.604	
	Control t-1 'Other'	0.997	0.997	0.207 – 4.810	
	Value t-1 'Negative'	0.001	0.144	0.048 – 0.438	
	Value t-1 'None'	0.089	0.214	0.036 – 1.266	

Table 7.19 Predictors selected for *value* using Binary Logistic Regression

$$\begin{aligned}\ln\left(\frac{p_{negative}}{1-p_{negative}}\right) &= -2.514 + 1.832 type\_of\_outcome_{\text{BlackOut}} - 0.114 type\_of\_outcome_{\text{DistanceTooFar}} \\ &+ 2.125 type\_of\_outcome_{Distance TooShort} + 0.924 type\_of\_outcome_{PositiveAcceleration} - 0.003 control\_t\_1_{other} \\ &- 16.414 type\_of\_outcome_{\text{Re}Started} + 2.592 type\_of\_outcome_{Success} + 1.570 control\_t\_1_{irrelevant} \\ &+ 2.425 publishing\_outcome_{Published} - 0.339 control\_t\_1_{high} + 1.375 control\_t\_1_{low} - 1.540 control\_t\_1_{none} \\ &- 0.112 control\_t\_1_{medium} - 1.936 control\_t\_1_{negative}\end{aligned} \quad \text{Eq. 7.15}$$

$$\begin{aligned}\ln\left(\frac{p_{irrelevant}}{1-p_{irrelevant}}\right) &= -0.924 - 1.103 publishing\_outcome_{Published} + 0.815 control\_t\_1_{high} \\ &- 1.126 average\_quality\_tutoring\_feedback_{(\text{-inf-}1.53]} + 0.625 control\_t\_1_{other} + 1.264 control\_t\_1_{low} \\ &+ 0.332 average\_quality\_tutoring\_feedback_{(1.53\text{-}1.595]} + 1.543 control\_t\_1_{irrelevant} + 2.201 control\_t\_1_{medium}\end{aligned} \quad \text{Eq. 7.16}$$

$$\begin{aligned}\ln\left(\frac{p_{other}}{1-p_{other}}\right) &= -0.630 - 1.983 publishing\_outcome_{Published} + 1.522 control\_t\_1_{high} \\ &- 0.030 average\_quality\_tutoring\_feedback_{(\text{-inf-}1.53]} + 2.416 control\_t\_1_{medium} + 1.451 control\_t\_1_{low} \\ &- 0.177 average\_quality\_tutoring\_feedback_{(1.53\text{-}1.595]} + 0.466 control\_t\_1_{irrelevant} + 2.857 control\_t\_1_{other}\end{aligned} \quad \text{Eq. 7.17}$$

Table 7.21 shows that the *retrospective outcome emotions* network has difficulty in recognising emotions that depend on identifying 'Positive' *value,* e.g. *value* has a sensitivity of 0.433 in Table 7.22, i.e. only 43.3% of the cases with 'Positive' *value* are classified correctly. As a result, gratitude is not identified and joy and pride are identified with probabilities with sensitivities of 0.100 and 0.300 respectively. This same network shows an improved reasoning for emotions that depend on identifying 'Negative' *value* ('Negative' value has a sensitivity of 0.945 in Table 7.22), such as anger, sadness and shame. However, the sensitivity achieved with these emotions also depends on how the *retrospective outcome emotions*

network classifies *control*. As a result, anger, sadness and shame are identified with probabilities of sensitivity of 0.770, 0.450 and 0.400 respectively. As a result, anger is the emotion that it is classified with more accuracy and precision.

Dependent variables (DVs)	*Predictors (IVs)*	*Sig. (p-value)*	*Odds ratios*	*95% Confidence Intervals (C.I.)*	*% cases correctly classified*
Control 'Irrelevant'	Publishing outcome 'Published'	0.022	0.332	0.129 – 0.854	57.9 (150 cases)
	Average quality tutoring feedback '(-inf-1.53]'	0.154	0.324	0.069 – 1.525	
	Average quality tutoring feedback '(1.53-1.595]'	0.567	1.394	0.447 – 4.345	
	Control t-1 'High'	0.166	2.258	0.714 – 7.147	
	Control t-1 'Irrelevant'	0.021	4.676	1.259 – 17.369	
	Control t-1 'Low'	0.039	3.540	1.066 – 11.751	
	Control t-1 'Medium'	0.068	9.037	0.851 – 95.976	
	Control t-1 'Other'	0.443	1.868	0.378 – 9.235	
Control 'Other'	Publishing outcome 'Published'	8.57E-5	0.138	0.051 – 0.370	
	Average quality tutoring feedback '(-inf-1.53]'	0.960	0.970	0.296 – 3.177	
	Average quality tutoring feedback '(1.53-1.595]'	0.740	0.838	0.296 – 2.375	
	Control t-1 'High'	0.007	4.580	1.516 – 13.840	
	Control t-1 'Irrelevant'	0.540	1.594	0.359 – 7.071	
	Control t-1 'Low'	0.015	4.268	1.319 – 13.808	
	Control t-1 'Medium'	0.045	11.202	1.053 – 119.119	
	Control t-1 'Other'	1.62E-5	17.414	4.750 – 63.833	

Table 7.20 Predictors selected for *control* using Multinomial Logistic Regression

Retrospective outcome emotions										
Ob-served	Predicted									
	Anger	Gratitude	Joy	Pride	Sadness	Shame	Spec-ificity	Sensi-tivity	Pre-cision	Network accuracy
Anger	77	3	1	1	10	8	0.613	0.770	0.554	0.504
Gratitude	5	0	0	1	3	1	0.932	0.000	0.000	
Joy	11	1	2	3	3	0	0.992	0.100	0.500	
Pride	11	8	1	9	0	1	0.974	0.300	0.600	
Sadness	22	3	0	0	27	8	0.880	0.450	0.529	
Shame	13	2	0	1	8	16	0.918	0.400	0.471	

Table 7.21 Confusion matrix corresponding to *retrospective outcome emotions*

To assess the agreement between student self-reports and the predictions of *retrospective* outcome emotions network, we employ once more Cohen's Kappa (Landis and Koch 1977, Cohen 1960). Accordingly, we also state our hypotheses as:

H_0- The degree of agreement between the student self-report and *PlayPhysics' retrospective emotions* network classification is random (κ=0)

H_A- There is a reasonable degree of agreement between the student self-report and PlayPhysics' *retrospective outcome emotions* network classification

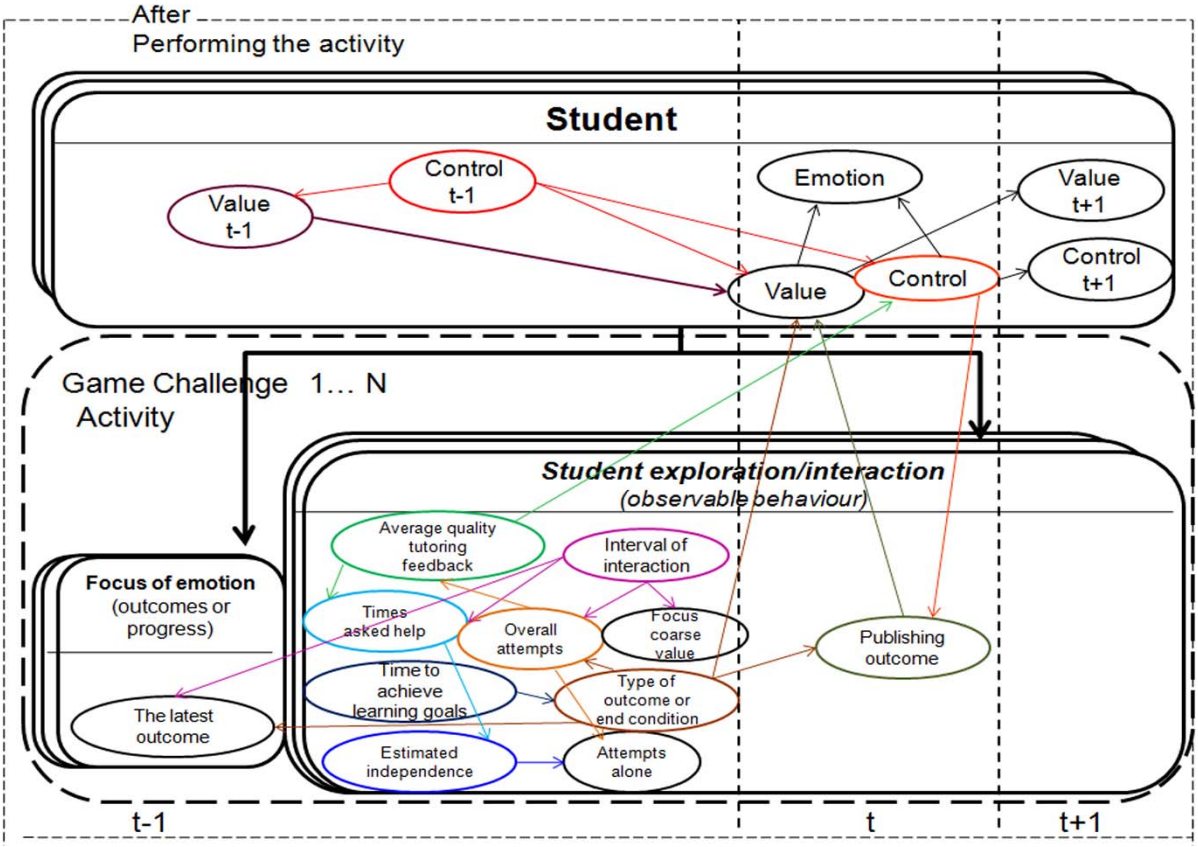


Figure 7.6 *Retrospective outcome achievement* emotions network

Value of retrospective outcome emotions						
Observed	Predicted					
	Negative	Positive	Specificity	Sensitivity	Precision	Network accuracy
Negative	189	11	0.433	0.945	0.847	0.827
Positive	34	26	0.945	0.433	0.703	

Table 7.22 Confusion matrix corresponding to *value*

Table 7.24 presents the results of Cohen's Kappa for *control*, *value* and emotion for 260 cases in the test data set. Along with the criteria of agreement for Kappa by Landis and Koch (1977), Cohen's Kappa shows a 'fair agreement' for *emotion* and *control*, i.e. κ=0.310; p<0.01 and κ=0.306; p<0.01 respectively. In this case, Cohen's Kappa shows a 'moderate agreement' with *value*, i.e. κ=0.437; p<0.01. As a result, we reject the null hypothesis and it may also be said that the results obtained were not random, since the *retrospective outcome emotions* network has a fair to moderate accuracy of classification.

Control of *retrospective outcome* emotions							
Observed	Predicted						
	Irrelevant	Other	Self	Specificity	Sensitivity	Precision	Network accuracy
Irrelevant	32	37	11	0.872	0.4	0.582	0.562
Other	14	85	11	0.54	0.773	0.552	
Self	9	32	29	0.884	0.414	0.569	

Table 7.23 Confusion matrix corresponding to *control*

Cohen's Kappa for the retrospective outcome achievement emotions network		
Dependent variable	*K*	*Significance*
Emotion	0.310	4.377E-21
Value	0.437	1.881E-13
Control	0.306	1.127E-12

Table 7 24 Cohen's Kappa for the *retrospective outcome emotions* network

This section presented a comprehensive explanation of the approach followed to define and assess the PlayPhysics' emotional student model including only observable behaviour variables related to a GBL environment context. Results corresponding to applying Cohen's Kappa for assessing the agreement between student self-reports and prospective outcome, activity and retrospective outcome emotions networks were presented. Finally, it can be said that PlayPhysics' emotional student model using contextual variables achieves a fair to moderate accuracy, validating the alternative hypotheses. However, it has potential for enhancement, since the values obtained for Cohen's Kappa are in a range between 0.3 and 0.6.

7.3 Evaluation of emotional model with physiological variables

For conducting the experimentation of student interaction with PlayPhysics GBL environment using the Bluetooth Galvanic Skin Response (GSR), we have eight students who volunteered and interacted with PlayPhysics GBL environment for approximately fifteen or twenty minutes each. As a result, there was not enough data to define and evaluate the networks corresponding to the prospective outcome and retrospective outcome achievement emotions networks. Hence, we decided to evaluate whether the acquisition of students' GSR signal can enhance the accuracy of our model by defining and evaluating the activity emotions network. In total we have 46 instances available in order to achieve this goal. In this dataset, 21.7% of the cases correspond to anger, 37% to enjoyment, 19.6% correspond to boredom and 21.7% to frustration. All the contextual variables that were not constants were included and divided into two categories, as we did previously for deriving the activity emotions network. The raw value of the GSR signal corresponding to the timestamp in the interaction log was retrieved for each instance. The value of the GSR signal was converted into discrete,

two and three categories, in order to identify the division that best suits the classification. The final dataset employed is shown in Table 7.25. The conversion to discrete was performed with the unsupervised filter of WEKA 'Discretize', using equal frequency binning.

Predictor	*Category*	*N° cases*	*Predictor*	*Category*	*N° cases*
Times asked help	(-inf-0.5]	40	Raw value GSR signal (3 cat)	(-inf-755.5]	15
	(0.5-inf)	6		(755.5-791.5]	16
Attempts alone	(-inf-0.5]	15		(791.5-inf)	15
	(0.5-inf)	31	Value t-1	Negative	25
Estimated independence	(-inf-1.5]	16		None	7
	(1.5-inf)	30		Positive	14
Overall attempts	(-inf-0.5]	15	Control t-1	High	19
	(0.5-inf)	31		Irrelevant	2
Interval of interaction	(-inf-380.5s]	23		Low	16
	(380.5s-inf)'	23		Medium	6
Focus coarse value	(-inf-3.09]	23		Self	3
	(3.09-inf)	23			
Raw value GSR signal (2 cat)	(-inf-772.5]	23			
	(772.5-inf)	23			

Table 7.25 Potential predictors of *activity emotions including GSR signal*

To gain more insight into the variables that should be selected, we applied binary logistic regression for control (categories 'High' and 'Low') and multinomial logistic regression for value (categories 'Negative', 'Positive' and 'None') including and excluding the GSR signal. Results are shown in Tables 7.26 to 7.27. As can be observed, the result was the same including and excluding the random variable GSR. Therefore, the raw value of the GSR signal was not selected using logistic regression, in conjunction with the forward conditional or stepwise procedure (discussed in Chapter 3) neither for value nor control. However, from the Pearson correlations it is observed that the GSR signal is associated to various contextual variables, and it may be that these interactions can increase the accuracy of classification of the network. However, it is noted that dividing the GSR into two categories may give a better chance of being related to student value ($r(46) = 0.267$; $p>0.05$). As a result we employed the dataset of Table 7.5 and excluding or including only the raw values of the GSR signal divided into two categories to create the activity emotions networks show in Figures 7.7 and 7.8. The structure of these networks was created applying the Necessary-Path Condition learning algorithm in combination with Pearson correlations and the results of the Binary and Multinomial Logistic Regression (BLR/MLR). EM learning algorithm was employed to learn the CPTs (see Appendix K). The logistic regression equation 7.18 shows the logit of having a student with 'high' control (p_{high}) referenced against the 'low' control group. Equations 7.19 to 7.20

show the logit of having a student with 'negative' value ($p_{negative}$) and a student with 'none' value (p_{none}) respectively, both referenced against the group 'positive' value.

The networks generated using the data sets including and excluding the GSR raw value are shown in Figures 7.7 and 7.8 respectively. It can be observed that the GSR signal is not associated to control nor value in Figure 7.8. Furthermore, there is no association between the predictors selected for control and value and the raw value of the GSR signal. However, in parallel a sub-network, shown in Figure 7.9, was created that included the raw value of the GSR signal, the focus coarse value and the number of times that the student asked for help. It was observed that the latter was linked to value in the previously derived activity emotions network (see Figure 7.5), so it may be that if we had more participants and collected more data in future experimentation, this variable may demonstrate a stronger relationship with control or value or one of their selected predictors. However, it is observed that the raw value of the GSR signal is linked to student concentration and number of times that students asked for help. The latter is related to student effort. When the Pearson values of these associations are reviewed, $r(46) = -0.387$; $p<0.01$ and $r(46) = -0.487$; $p<0.01$ for number of times help was asked and focus coarse value respectively. It is noted that these are associations of medium strength, so between 9% and 25% of the variance is shared between these variables.

Dependent variables (DVs)	*Predictors (IVs)*	*Sig. (p-values)*	*Odds ratios*	*95% Confidence Intervals (C.I.)*	*% cases correctly classified*
Control 'High'	Value t-1 'Negative'	0.167	2.881	0.642 – 12.926	65.2 (30 cases)
	Value t-1 'None'	0.037	9.167	1.147 – 73.239	

Table 7.26 Predictors selected for *control* including/excluding GSR using BLR

Dependent variables (DVs)	*Predictors (IVs)*	*Sig. (p-value)*	*Odds ratios*	*95% Confidence Intervals (C.I.)*	*% cases correctly classified*
Value 'Negative'	Value t-1 'Negative'	0.003	31.651	3.173 – 315.725	76.1 (35 cases)
	Value t-1 'None'	0.045	30.109	1.087 – 833.916	
	Attempts alone '(-inf-0.5]'	0.246	0.275	0.031 – 2.433	
Value 'None'	Value t-1 'Negative'	0.187	4.675	0.474 – 46.136	
	Value t-1 'None'	0.052	15.320	0.979 – 239.798	
	Attempts alone '(-inf-0.5]'	0.117	5.527	0.651–46.904	

Table 7.27 Predictors selected for *value including/*excluding GSR using MLR

$$\ln\left(p_{high} \big/ 1 - p_{high} \right) = -1.299 + 1.058 value\_t-1_{negative} + 2.216 value\_t-1_{none} \qquad \text{Eq. 7.18}$$

$$\ln\left(\frac{p_{negative}}{1-p_{negative}}\right) = -2.065 + 3.455 value\_t-1_{negative} + 3.405 value\_t-1_{none} - 1.291 attempts\_alone_{(-\inf -0.5]} \quad \text{Eq. 7.19}$$

$$\ln\left(\frac{p_{none}}{1-p_{none}}\right) = -2.722 + 1.542 value\_t-1_{negative} + 2.729 value\_t-1_{none} + 1.710 attempts\_alone_{(-\inf -0.5]} \quad \text{Eq. 7.20}$$

To evaluate the accuracy of classification of the activity emotions networks excluding and including the GSR signal derived, stratified10-fold cross-validation was applied. Results corresponding to 50 test instances are shown in Tables 7.28 to 7.29. It was observed that in some network structures, which are the result of applying cross-validation, value was related to times asked help. The latter was simultaneously associated to the raw value of the GSR signal. Furthermore, the raw value of the GSR signal was sometimes associated to control t-1 or value t-1. When the results obtained in Tables 7.28 to 7.30 were compared with the results obtained in Tables 7.31 to 7.33, it was observed that by including the raw value of the GSR signal in combination with contextual variables the model increases its accuracy for reasoning about emotion from 0.532 to 0.700. Enjoyment is the emotion that is classified with the highest sensitivity and precision. However, it is observed that anger has a probability of sensitivity equal to 0.700, although its precision is not as high as that of boredom, which has a probability of sensitivity of 0.500 and precision of 0.714. It is also noted that the overall accuracy of classification for value is enhanced from 0.600 to 0.780. However, the overall accuracy of classification for control is slightly decreased from 0.840 to 0.760.

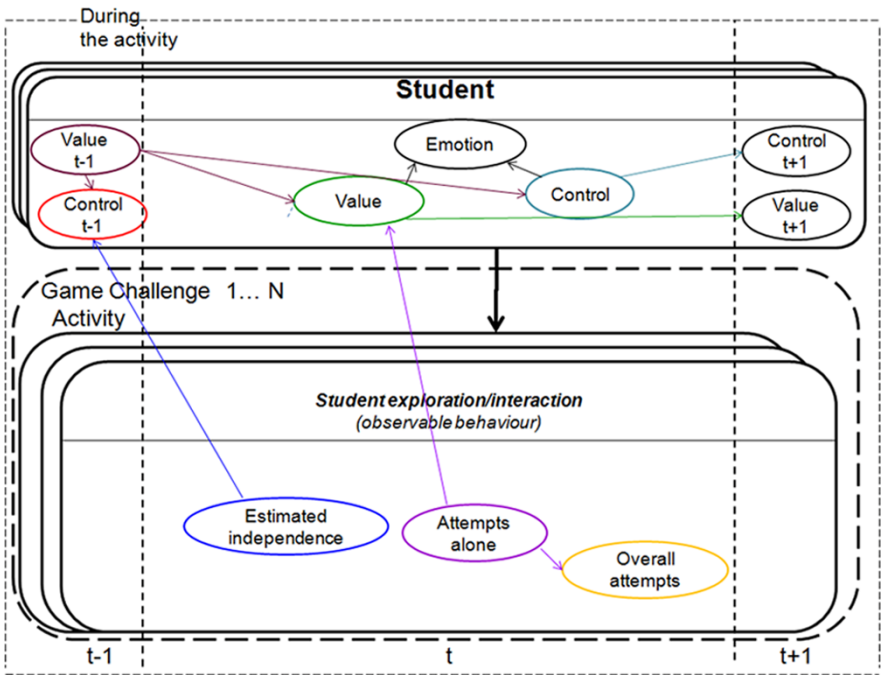


Figure 7.7 Activity emotions network generated with dataset excluding the GSR signal

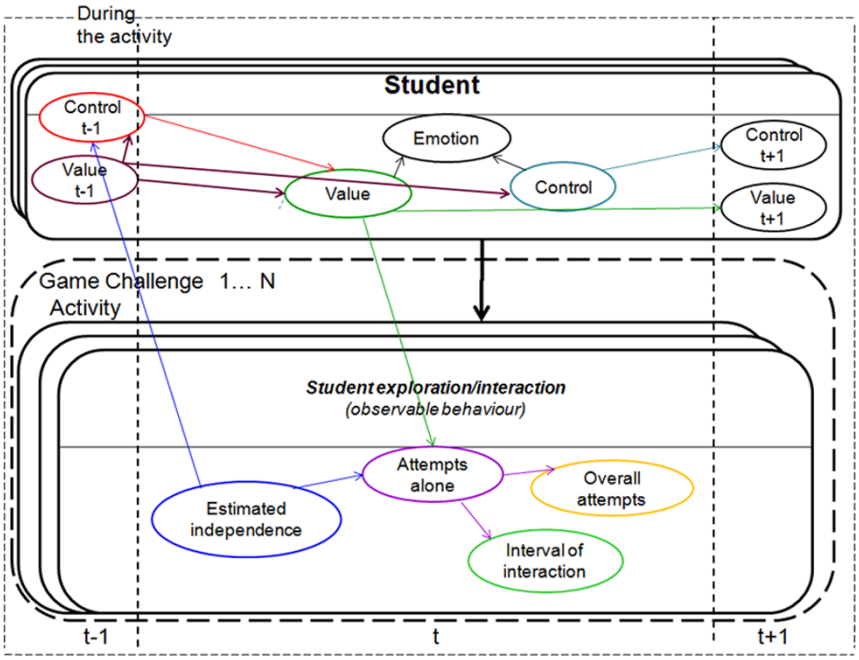


Figure 7.8 Activity emotions network generated with dataset including the GSR signal

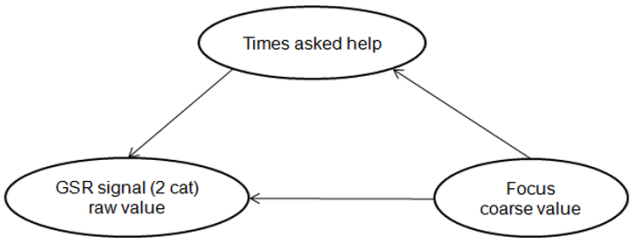


Figure 7.9 Sub-network created including the GSR signal

Activity emotions								
Observed	Predicted							
	Anger	Boredom	Enjoyment	Frustration	Specificity	Sensitivity	Precision	Network accuracy
Anger	8	0	1	1	0.625	0.800	0.348	0.520
Boredom	0	4	3	3	0.975	0.400	0.800	
Enjoyment	11	0	9	0	0.867	0.450	0.692	
Frustration	4	1	0	5	0.900	0.500	0.556	

Table 7.28 Confusion matrix corresponding to *activity emotions* excluding *GSR*

Value of *activity* emotions							
Observed	Predicted						
	Negative	None	Positive	Specificity	Sensitivity	Precision	Network accuracy
Negative	17	0	3	0.567	0.850	0.567	0.600
None	2	4	4	1.000	0.400	1.000	
Positive	11	0	9	0.767	0.45	0.563	

Table 7.29 Confusion matrix corresponding to *value* with data set excluding *GSR*

Control of *activity* emotions						
Observed	Predicted					
	High	Low	Specificity	Sensitivity	Precision	Network accuracy
High	29	1	0.650	0.967	0.810	0.840
Low	7	13	0.967	0.650	0.929	

Table 7.30 Confusion matrix corresponding to *control* with data set excluding *GSR*

Activity emotions								
Observed	Predicted							
	Anger	Boredom	Enjoyment	Frustration	Specificity	Sensitivity	Precision	Network accuracy
Anger	7	0	1	2	0.875	0.700	0.583	0.700
Boredom	1	5	3	1	0.950	0.500	0.714	
Enjoyment	0	0	19	1	0.866	0.950	0.826	
Frustration	4	2	0	4	0.900	0.400	0.500	

Table 7.31 Confusion matrix corresponding to *activity emotions* including *GSR*

Value of *activity* emotions							
Observed	Predicted						
	Negative	None	Positive	Specificity	Sensitivity	Precision	Network accuracy
Negative	16	3	1	0.900	0.800	0.842	0.780
None	2	4	4	0.925	0.400	0.571	
Positive	1	0	19	0.833	0.950	0.792	

Table 7.32 Confusion matrix corresponding to *value* with data set including *GSR*

Control of *activity* emotions						
Observed	Predicted					
	High	Low	Specificity	Sensitivity	Precision	Network accuracy
High	24	6	0.700	0.800	0.800	0.760
Low	6	14	0.800	0.700	0.700	

Table 7.33 Confusion matrix corresponding to *control* with data set including *GSR*

For assessing the agreement between student self-reports and the predictions of the activity emotions network created using a dataset including the raw value of the GSR signal, we employ again Cohen's Kappa (Landis and Koch 1977, Cohen 1960). Our hypotheses are:

H_0 - The degree of agreement between the student self-report and *PlayPhysics' activity emotions network predictions,* which includes the GSR signal, is random (κ=0)

H_A - There is a reasonable degree of agreement between the student self-report and PlayPhysics' *activity emotions* network predictions including the GSR signal

Tables 7.34 and 7.35 show the results of Cohen's Kappa for control, value and emotion for 50 test instances excluding and including the GSR raw value respectively. Along with the criteria of agreement for Kappa by Landis and Koch (1977), see Table 7.9. Cohen's Kappa

shows a ‘moderate agreement’ for emotion and control, i.e. κ=0.576; p<0.01 and κ=0.500; p<0.010 respectively. In this case, Cohen’s Kappa shows a substantial agreement’ with value, i.e. κ=0.650; p<0.010. As a result, we can reject the null hypothesis and it can also be said that the results obtained were not random, since the activity emotions network including the raw value of the GSR signal shows a moderate to substantial accuracy of classification. When the results in Table 7.34 (excluding the GSR signal) are compared with the results in Table 7.35 (including the GSR signal) it is observed that the classification of emotion is elevated from fair agreement (0.358) to moderate agreement (0.576). The classification of value also improves from fair agreement (0.351) to substantial agreement (0.650) when including the GSR signal. However, the classification of control drops from substantial agreement (0.649) to moderate agreement (0.500) when including the GSR signal. Therefore, it may be stated that the GSR raw value enhances the classification of value rather than control. Thus, including the GSR raw value improves the overall classification of student activity emotions.

Cohen’s Kappa for the activity achievement emotions network		
Dependent variable	*K*	*Significance*
Emotion	0.358	3.41E-6
Value	0.351	5.39E-4
Control	0.649	1.96E-6

Table 7.34 Cohen's Kappa for the activity emotions network excluding GSR

Cohen’s Kappa for the activity achievement emotions network		
Dependent variable	*K*	*Significance*
Emotion	0.576	6.88E-12
Value	0.650	4.21E-10
Control	0.500	4.07E-4

Table 7.35 Cohen's Kappa for the activity emotions network including GSR

7.4 PlayPhysics GBL environment evaluation

To achieve further insight into the factors or features of PlayPhysics GBL environment that influenced student achievement emotions during interaction, we conducted a quantitative and qualitative evaluation of PlayPhysics GBL environment. The objective is to identify other features or factors that may enhance the classification of our emotional student model in the future.

7.4.1 Quantitative evaluation

As mentioned earlier in section 7.2.1, 118 students in the control and focus groups are used to perform the quantitative evaluation of PlayPhysics GBL environment. Students are all be-

tween 18 and 23 years old and are taking an introductory Physics module. Students in the focus and control groups solved a pre-test including concepts of physics taught by PlayPhysics GBL environment. Afterwards students in the focus group proceeded to interact with the PlayPhysics GBL environment, while students in the control group reviewed a video comprising a PowerPoint presentation covering the same concepts. Finally, students in both groups solved a post-test. As a result, this experiment can be viewed as an independent samples test, where the absolute and relative learning gains of students in both control and focus groups are compared. The absolute and relative learning gains are obtained as presented in Equations 7.21 to 7.22 respectively. This were defined by Hake (1998) and successfully applied in Muñoz et al. (2009b).

$$G = PostTest - \mathrm{Pr}\,eTest \qquad \text{Eq. 7.21}$$

$$G_{rel} = \frac{(PostTest - \mathrm{Pr}\,eTest)}{(100 - \mathrm{Pr}\,eTest)} \qquad \text{Eq. 7.22}$$

In this case, we want to compare the absolute and relative learning gains between the focus and control group samples and identify whether the difference between the two means is significant. To achieve this we applied the *t test*, since we are working with continuous data, e.g. absolute and relative learning gains. The *t test* is a parametric test (Kinnear and Gray 2010). Therefore, both a normal distribution of the population and equal variances are assumed. The data is examined using the Statistical Package for Social Science (SPSS) (IBM 2012). The summary of this analysis is shown in Table 7.36. As can be observed, when the absolute and relative learning gains were calculated there were cases where the relative learning gain was undefined, since some students obtained a 100% score on the pre-test. These cases are indicated under the column entitled “Missing” and the resultant valid cases are indicated under the column entitled “Valid” in Table 7.36. Since the resultant valid cases are not in equal number for both populations, the populations were balanced via the use of random sampling. The latter goal was achieved applying the Resample unsupervised filter in WEKA, which produces a random sub-sample of the dataset - in this case without replacement.

Both populations were thus reduced to 52 cases each. However, the data had to be re-examined in order to search for outliers that might bias the means of both groups. Figures 7.10 and 7.11 show the boxplots corresponding to the absolute and relative learning gains for the control and focus groups. An outlier is indicated by a circle, one of which is shown 1.5 box lengths above the box in Figure 7.10, and an extreme case is indicated by a star, which is 3 box lengths below the box. The number accompanying the sign identifies the case. Therefore, cases 65 and 97 were removed from the sample.

Case Processing Summary							
	Experimental group	Cases					
		Valid		Missing		Total	
		N	%	N	%	N	%
Absolute learning gain (G)	Control	72	61.0	46	39.0	118	100.0
	Focus	52	44.1	66	55.9	118	100.0
Relative learning gain (G_{rel})	Control	72	61.0	46	39.0	118	100.0
	Focus	52	44.1	66	55.9	118	100.0

Table 7.36 Data descriptive corresponding to the focus and control samples

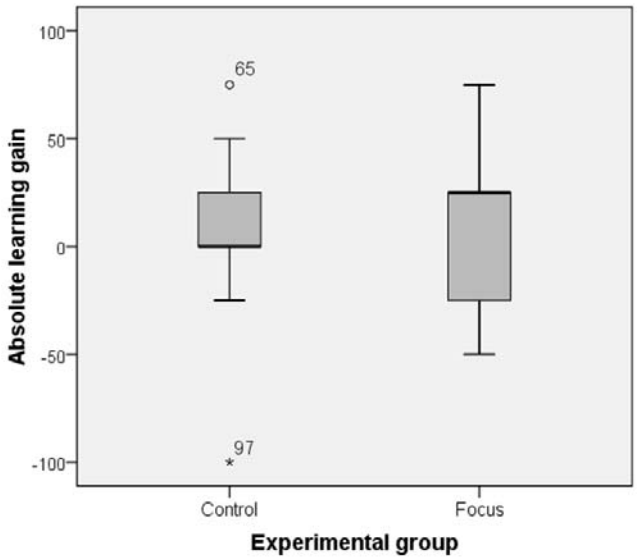


Figure 7.10 Boxplot corresponding to the absolute learning gain

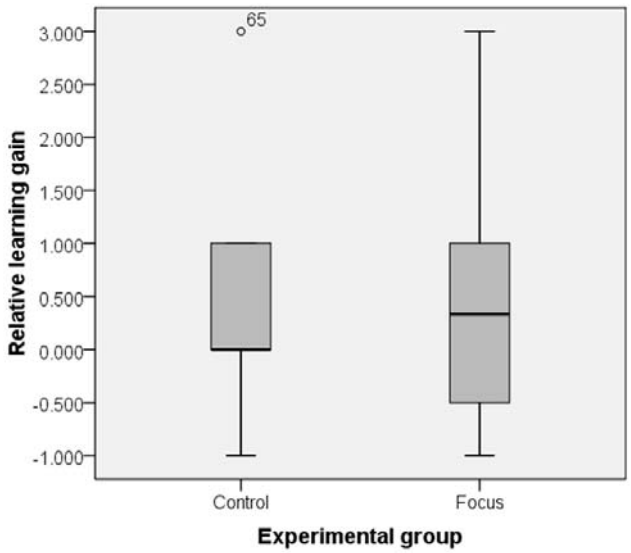


Figure 7.11 Boxplot corresponding to the relative learning gain

There were 50 cases from the control group and 52 cases from the focus group in the final dataset (see Appendix K). The mean and the standard deviation for the learning gains of the control and focus groups are shown in Table 7.37. Note that the means of the absolute and relative learning gains corresponding to the focus and control groups are different and the means of the focus group are slightly larger. At this moment, however, it cannot be said that the difference is significant - the *t-test* needs to be applied.

Experimental group	Learning gains	N	Mean	Std. Deviation	Standard Error Mean
Control group	Absolute learning gain	50	8.000	21.093	2.9836
	Relative learning gain		0.237	0.680	0.0969
Focus group	Absolute learning gain	52	9.620	32.897	4.562
	Relative learning gain		0.405	1.043	0.145

Table 7.37 Mean and standard deviations for the learning gains

The hypothesis can be stated in the following manner:

H_0 - Students in the focus and control groups have the same learning gain ($\mu_{control}=\mu_{focus}$).

H_A - Students in the control and focus groups have a different learning gain ($\mu_{control}<\mu_{focus}$) or ($\mu_{control}>\mu_{focus}$).

The results of the *t-test* are presented in Table 7.38, one row shows the results of the absolute learning gain and another shows the results of the relative learning gain. There are two values for *t* in each case. The upper *t* value corresponds to a traditional *t-test* with the pooled variance estimate, i.e. equal variances are assumed. However, the lower *t* value corresponds to a *t-test* performed with the Behrens-Fisher statistic T (Kinnear and Gray 2010), where variance estimates are not pooled, i.e. equal variances not assumed. This is important, since the degrees of freedom (df) are different. The values corresponding to 'equal variances not assumed' are adjusted using the Welch-Satterthwaite formula. Levene's test is consulted to select the appropriate value of *t*.

The Levene's test (F) examines the assumption of homogeneity of variance. The variances are not homogeneous, since the p-value is smaller than 0.05. Therefore, the value of *t* in the lower row is employed for the absolute and relative learning gains. The value of *t* for the absolute learning gain can be stated as t(87.319) = -0.296 and the value of t for the relative absolute learning gain can be stated as t(88.148) = -0.971. The negative sign is due to the subtraction of the mean of the focus group from the mean of the control group. As can be observed the value of *t* for the absolute and relative learning gains are not significant at the 0.05 level, so we are obliged to accept the null hypothesis (H_0) that students in the focus and control group have the same learning gain.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Absolute learning gain	Equal variances assumed	15.077	0.000	-0.294	100	0.769	-1.615	5.496	-12.519	9.288
	Equal variances not assumed			-0.296	87.319	0.768	-1.615	5.451	-12.449	9.218
Relative learning gain	Equal variances assumed	5.975	0.016	-0.963	100	0.338	-1.686E-1	0.175	-5.159E-1	0.179
	Equal variances not assumed			-0.971	88.148	0.334	-1.686E-1	0.174	-5.138E-1	0.176

Table 7.38 Results of the *t-test* corresponding to students' learning gains

However, it can be confirmed that students conducting independent learning with PlayPhysics have a positive absolute or relative learning gain - it does not matter that the students interacted either with the PowerPoint presentation in a video format or with the GBL environment. This can be confirmed statistically if we apply a *one-sample t-test*. The hypotheses can be stated as:

H_0 - Students do not have a learning gain ($\mu_G \leq 0$) or ($\mu_{Grel} \leq 0$)

H_A- Students have a different absolute or relative learning gain using PlayPhysics for conducting independent self-study ($\mu_G > 0$) or ($\mu_{Grel} > 0$)

The one-sample statistics and test results, corresponding to a population of 100 students, are shown in Tables 7.39 and 7.40 respectively. The cases corresponding to identifiers 65, 18, 22 and 30 were identified as outliers and removed (see Appendix K). In Table 7.39, it can be observed that the mean of the relative learning gain is approximately 0.23, whereas the mean of the absolute learning gain approximately 5.8 and their standard deviations are 0.77 and 27.5 respectively.

	N	Mean	Dtd. Deviation	Std. Error Mean
Relative learning gain (G_{rel})	100	0.229	0.767	0.077
Absolute learning gain (G)	100	5.750	27.491	2.749

Table 7.39 One-sample statistics of relative and absolute learning gains

From Table 7.40, it can be noted that the value of t for the relative learning gain and the absolute learning gain can be stated as t(99) = 2.989 and t(99) = 2.092. Also, it can be observed that this is a two-tailed test, since we are saying that absolute and relative learning gains may exist and may be negative. The critical values of *t* are significant at the 0.05 level (0.004 and 0.039 p-values for absolute and relative learning gains respectively). Therefore, we are obliged to accept the alternative hypothesis (H_A), which states that students that conduct self-study with PlayPhysics have a different absolute or relative learning gain.

	Test Value = 0					
					95% Confidence Interval of the Difference	
	T	Df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Relative learning gain	2.989	99	0.004	0.229	0.077	0.381
Absolute learning gain	2.092	99	0.039	5.750	0.300	11.200

Table 7.40 Results one-sample *t-test* of relative and absolute learning gains

7.4.2 Qualitative evaluation

From one-hundred and eighteen students in the focus group with complete data, eighty seven answered the qualitative questionnaire (Appendix H). This questionnaire was designed with the main purpose of determining participants' agreement with a set of statements related to PlayPhysics's features: fantasy, encouragement of curiosity, challenge, ease of interaction and appropriateness of the learning companion's behaviour. Answers to questions are designed as a Likert scale (Bertram 2012) comprising 5-points or values that range from 'Highly Disagree' to 'Highly Agree', where 'Highly Disagree' (HD) = 1 , 'Disagree' (D) = 2, 'Neither Agree or Disagree' (NAD) = 3, 'Agree' (A) = 4 and 'Highly Agree' (HA) = 5.

In this population of students, there were 33 females and 54 males, corresponding to approximately 38% and 62% of the student population respectively. When exploring student perceptions of PlayPhysics' fantasy and the appeal of the storytelling, see Figure 7.12, approximately 42.53% of students 'Agree' and 8.05% 'Highly Agree' that the story is appealing, while 10.34% 'Disagree' and 9.20% 'Highly Disagree'. Including game contexts other than a space adventure might encourage these students and alleviate their feelings of boredom, as they considered the space theme unattractive. Some students, however, found the game entertaining.

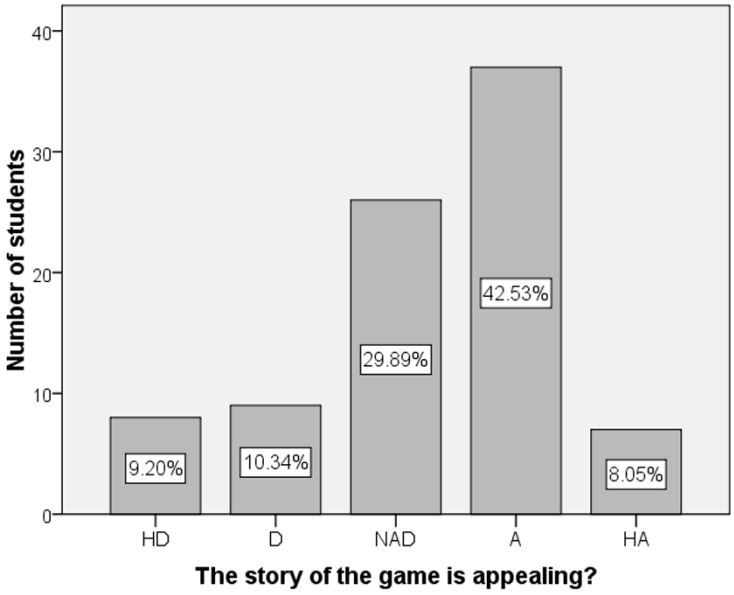


Figure 7.12 Student perception of PlayPhysics GBL environment storytelling

When exploring whether PlayPhysics GBL environment promotes student cognitive curiosity, students were asked whether PlayPhysics encouraged them to replay and increase their score (Figure 7.13) 42.53% 'Agreed' and 13.79% 'Highly Agreed' with this proposition, whereas 16.09% students 'Disagreed' and 10.34% 'Highly Disagreed'. Another question related to the encouragement of student cognitive curiosity is whether PlayPhysics GBL envi-

ronment encourages persistence and the solving of other game challenges, (see Figure 7.14). It was observed that approximately 39.08% students 'Agreed' and 10.34% 'Highly Agreed' with this proposition, whereas about 16.09% students 'Disagreed' and 9.20% 'Highly Disagreed'. It has been demonstrated that incorporating elements of complexity, novelty, surprise and incongruence can arouse cognitive curiosity (Berlyne 1960). For example, including game challenges where the student can select the complexity, instead of randomly assigning it may increase their curiosity and avoid boredom. One student suggested including more interaction variables in order to increase the complexity of the game challenge.

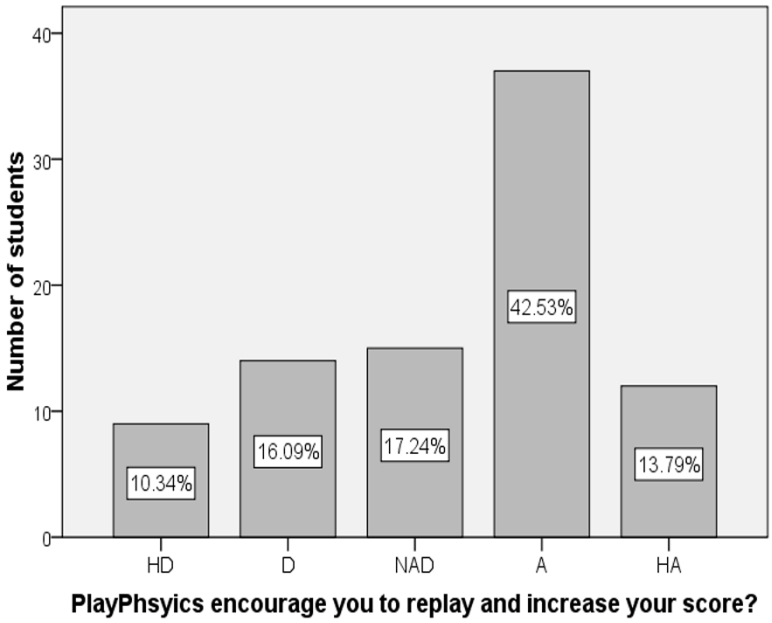


Figure 7.13 Student perception of PlayPhysics encouragement for replaying

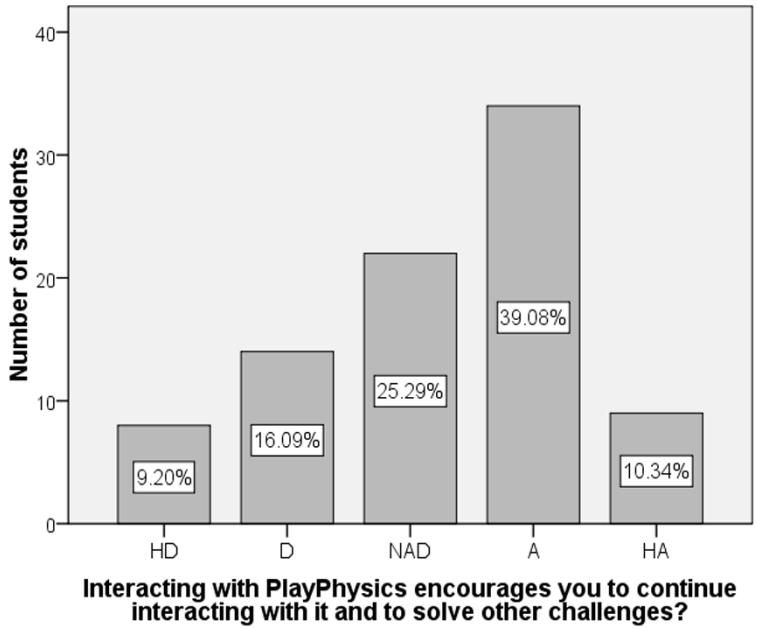


Figure 7.14 Student perception of PlayPhysics arousal for approaching challenges

To evaluate whether PlayPhysics' fantasy arouses students' cognitive curiosity, the questionnaire asked if students were encouraged to consider other situations where they might have to apply the same principles used to solve PlayPhysics' game challenge. Results are shown in Figure 7.15, around 43% of students 'Agreed' and 9.20% 'Highly Agreed' with this statement, whilst 16.09% 'Disagreed' and 10.34% 'Highly Disagreed'. As mentioned earlier, some students may benefit from a different game context to enable them to better grasp the concepts. Also, it may be necessary to present more than one challenge to the student involving the same topics in order to ensure that the student has really achieved the learning goals.

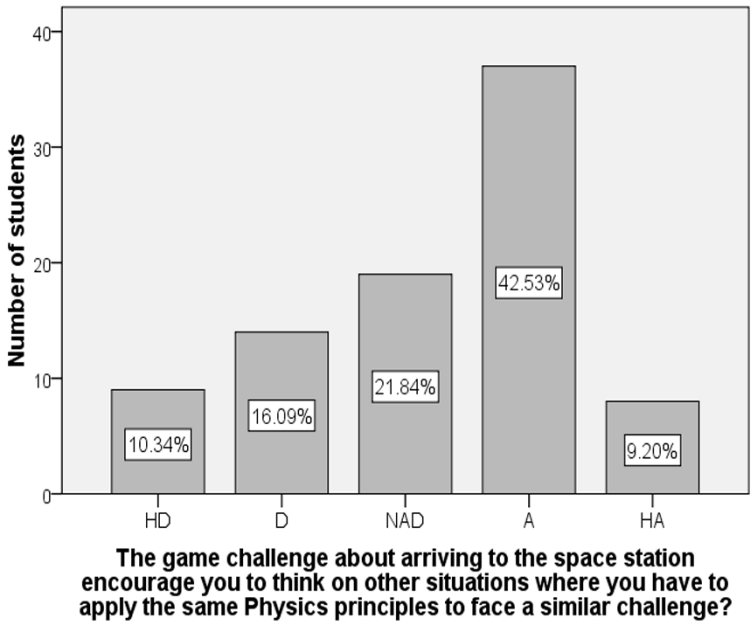


Figure 7.15 Student perception of PlayPhysics fantasy on students' cognitive curiosity

Regarding the clarity of the game challenge goals, students were asked whether they thought that the goal of the game challenge was precise. The bar graph in Figure 7.16 shows that 48.28% of students 'Agreed' and 14.94% of students 'Highly Agreed' with this statement, whilst 13.79% of students 'Disagreed' and 4.60% 'Highly Disagreed'. Most of students agree that what they have to do in order to successfully achieve the game challenge was clear enough. This may be the reason that some students reported enjoyment, since challenge has been shown to influence intrinsic motivation (Malone 1981), which in turn has shown to create a sense of efficacy whilst dealing with the environment. However, from the set of students, who believed that instructions were not clear enough, some suggested being more explicit in the game rules. For example, it should be made explicit that you cannot travel backwards, i.e. move in the direction away to Athena space station. A small amount of students also found a few phrases confusing. This may be owing to the fact students that are

not native English speakers but are interacting with an educational game in English that uses some terminology to which they are not accustomed. However, the majority of students have shown to have a proficient level of English.

It is also necessary to discover whether the feedback provided was appropriate, since students have shown to experience feelings of frustration when they do not have sensible feedback on the progress that they have made towards the learning outcomes (Hounsell 2009). As a result, students were asked whether the feedback provided through sound, graphics and the game score assisted them in achieving awareness of their progress. Results in Figure 7.17 show that 47.13% of students 'Agreed' and 6.90% 'Highly Agreed' with the appropriateness of the feedback provided by PlayPhysics GBL environment, while 12.64% 'Disagreed' and 10.34% 'Highly Disagreed'.

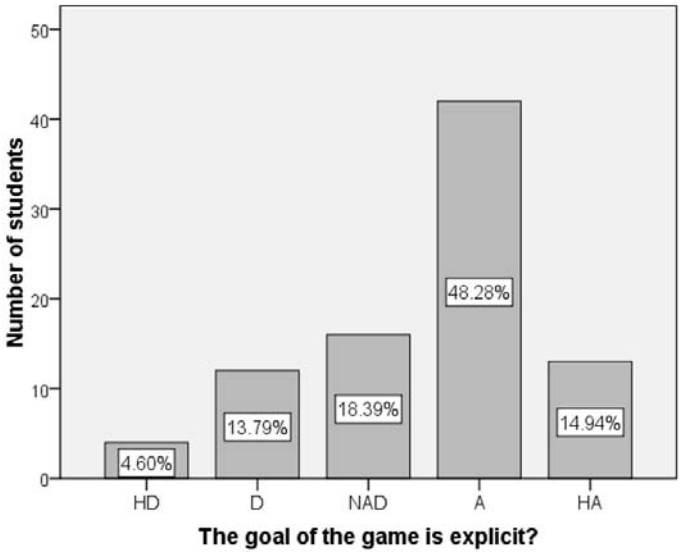


Figure 7.16 Evaluating the clarity of PlayPhysics game goals

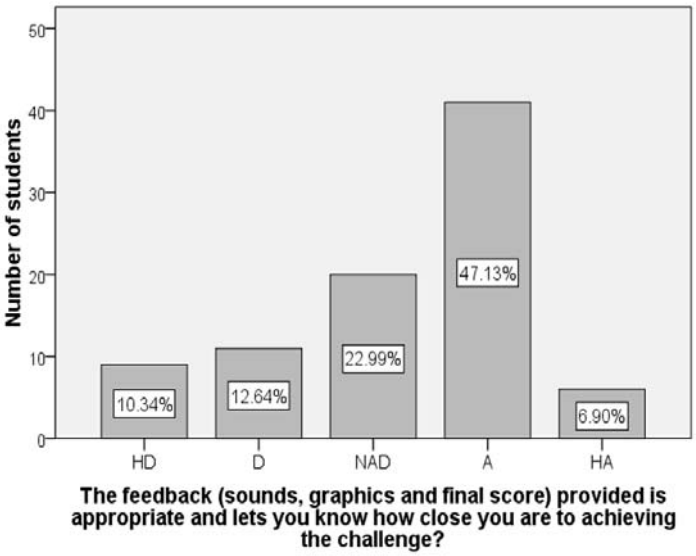


Figure 7.17 Appropriateness of PlayPhysics game feedback

Some of the feedback in PlayPhysics, e.g. more detailed clues, was provided by the M8 robot, PlayPhysics' learning companion. It was observed that 81 out of 118 students (about 69% of students) in the focus group asked for M8's help. However, M8 robot's main objective was to frequently remind students about reporting their emotional state. M8 provided some emotional response, such as mirroring students' emotions or encouraging students to keep interacting with PlayPhysics GBL environment. Students were asked about the appropriateness of M8's knowledge and feedback. In this specific case, students' opinions were similarly distributed, see Figure 7.18. Since 29.89% of students 'Agreed' and 4.60% 'Highly Agreed' with the appropriateness of M8's feedback, while 25.29% 'Disagreed and 11.49% 'Highly Disagreed'. Some students commented that M8 should give more than clues, so instead of simply reminding them about the relevant concepts, it should instead give the definition of those concepts. Students also thought that M8 should focus more on giving cognitive feedback than emotional feedback.

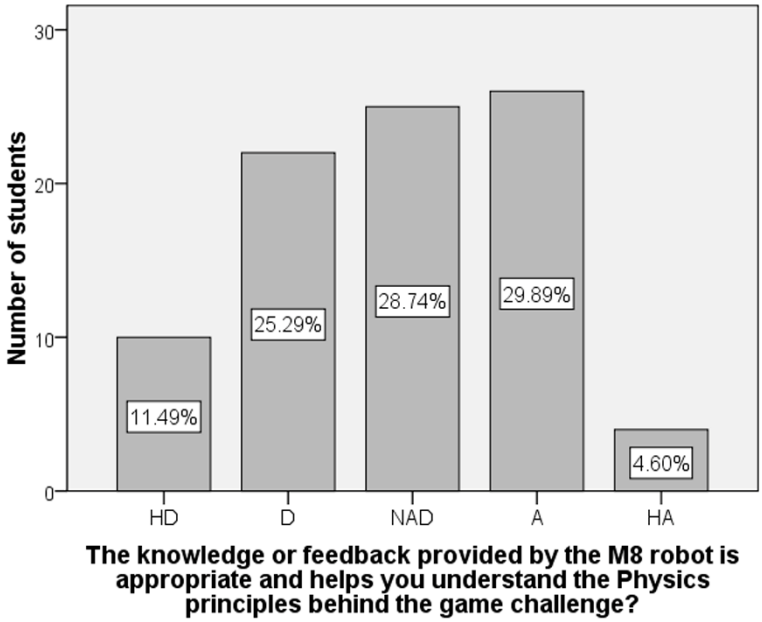


Figure 7.18 Evaluation of the appropriateness of M8 feedback

Students were asked about the achievability of PlayPhysics GBL environment game goal. The bar chart in Figure 7.19 shows that 49.43% of students 'Agreed' that it was achievable and 16.09% 'Highly Agreed', whereas 13.79% of the population of students 'Disagreed' and 6.90% 'Highly Disagreed' with this proposition. The agreement of the majority of the students with this statement may explain why some students reported feeling enjoyment while interacting with PlayPhysics GBL environment. However, students that disagreed with this statement may have found that the complexity of the challenge was higher than expected or that the available feedback was not appropriate to help clarify their misconceptions. Also, the complexity of the challenge could increase from the students' viewpoint if they struggled with

the English in the game. It may also be possible that they found PlayPhysics' GUI difficult to use. We asked students whether they think that PlayPhysics assisted them in enhancing their grasping of physics concepts and principles overall, see Figure 7.20. About 46% of students 'Agreed' and 11.49% 'Highly Agreed' with this proposition, whilst 13.79% 'Disagreed' and 8.05% 'Highly Disagreed'. These results are supported by the findings discussed in section 7.4.1, where it was found that students have an absolute and relative learning gain.

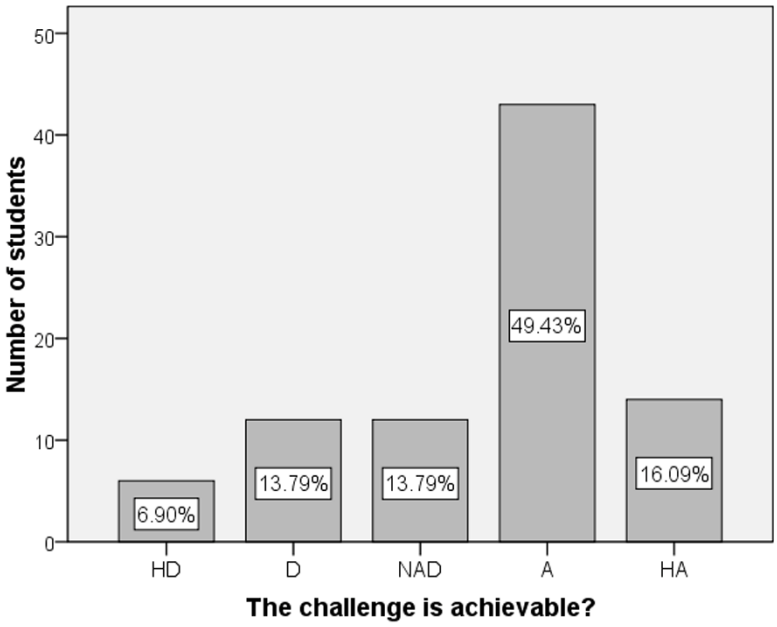


Figure 7.19 Student perceptions about the achievability of PlayPhysics game goals

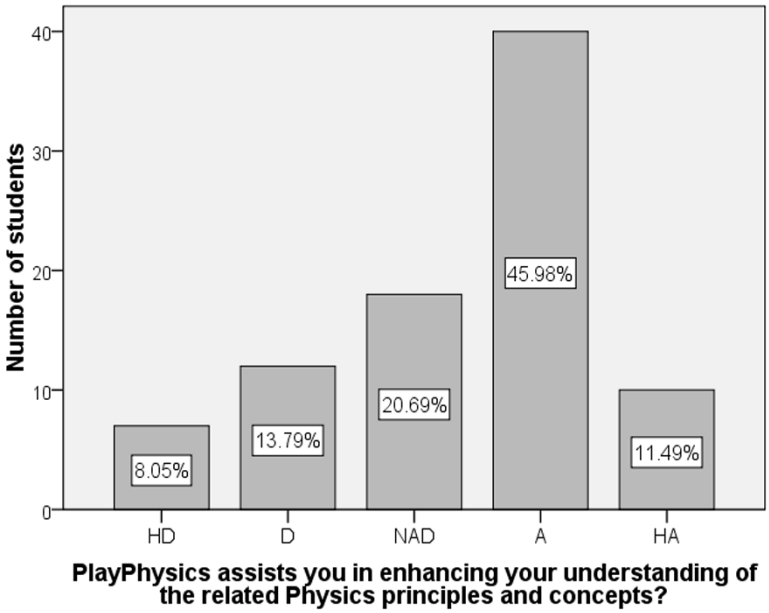


Figure 7.20 PlayPhysics enhances student grasp of physics concepts

To discover students' perceptions about PlayPhysics GBL environment ease of use, we asked them whether they believed that the GUI was intuitive. Results are shown in Figure

7.21, around 41% of students 'Agreed' with this statement, while 16.09% 'Disagreed' and 6.90% 'Highly Disagreed'. Some students mentioned that they felt very entertained while interacting with PlayPhysics GBL environment, so we assumed that for these students its GUI was easy to use and they enjoyed the experience. However, some students mentioned that instead of using arrows to set the velocity, it would be desirable to type the value of the velocity. Other students mentioned that a more detailed explanation of the controls is required. Some students also said that they found some words of the instructions given in English difficult to understand and as a result, they found instructions unclear. This may explain in some cases students' self-reports of frustration and anger.

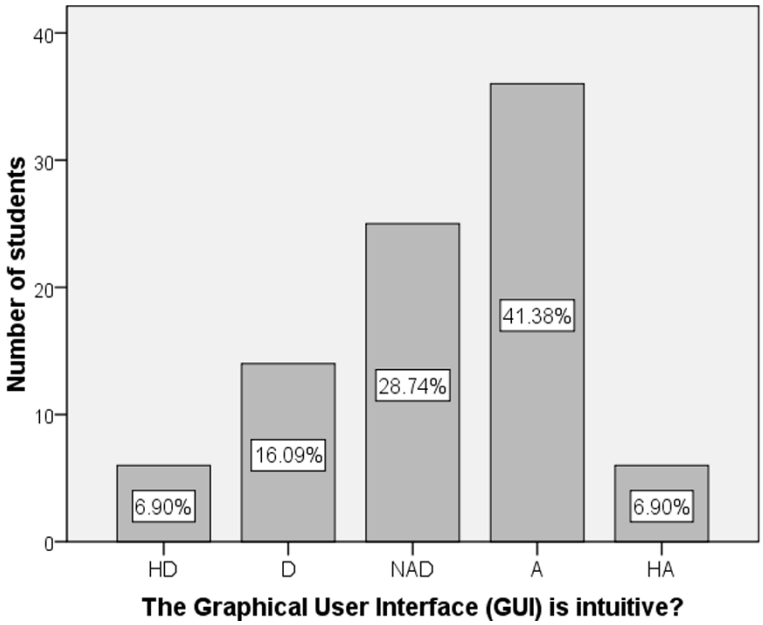


Figure 7.21 Evaluation of PlayPhysics GBL environment ease of use

Finally, we focus on evaluating the appropriateness of M8's emotional behaviour. In this research, M8 was limited to mirroring students' self-reported emotion, showing affinity or understanding and encouraging repeated attempts at the game challenge until the learning goals were achieved, as was explained in detail in Chapter 6. Figure 7.22 shows that 35.63% of students 'Agreed' and 8.05% 'Highly Agreed' that M8's emotional behaviour was appropriate. However, 13.79% of students 'Disagreed' and 9.20% 'Highly Disagreed' with this proposition. These results may help explain students' self-reports of feeling grateful or angry and frustrated while interacting with PlayPhysics GBL environment. One student mentioned that the M8 robot was annoying, since he/she perceived that was very insisting about not forgetting to self-report his/her emotion. This agrees with what Conati and Maclaren (2009) reported about using learning companions to remind students to self-report their emotional states. Other students considered that providing an emotional response is not as useful and valuable as the feedback provided about physics concepts.

As mentioned earlier, there are 54 males and 33 females in the student population that answered the qualitative questionnaire and we considered it important to compare female and male student opinions with respect to the appropriateness of M8 emotional behaviour, see Figure 7.23. This involves two independent samples of ordinal data. A non parametric test is used, since it cannot be assumed that this data has a normal distribution with the same variance. Therefore, to evaluate this question we employ the Mann-Whitney U test (Bertram 2012, Kinnear and Gray 2010), which is a non parametric alternative to the independent-samples t-test.

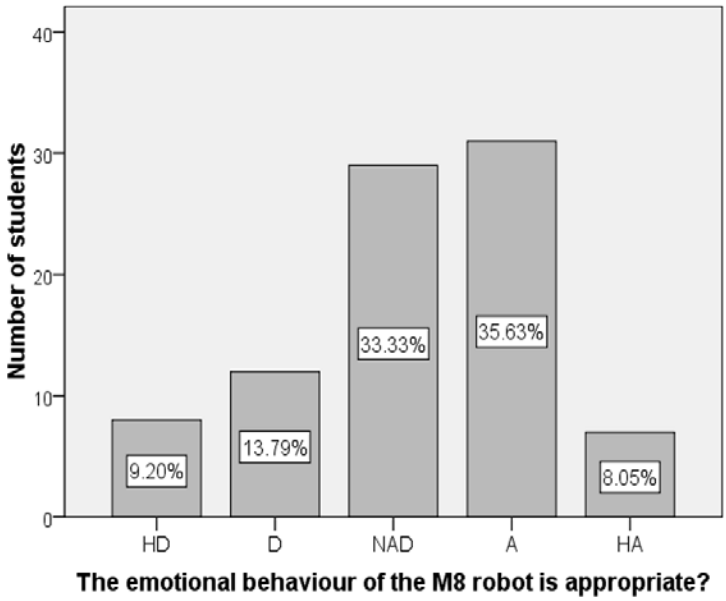


Figure 7.22 Evaluation of the appropriateness of M8's emotional behaviour

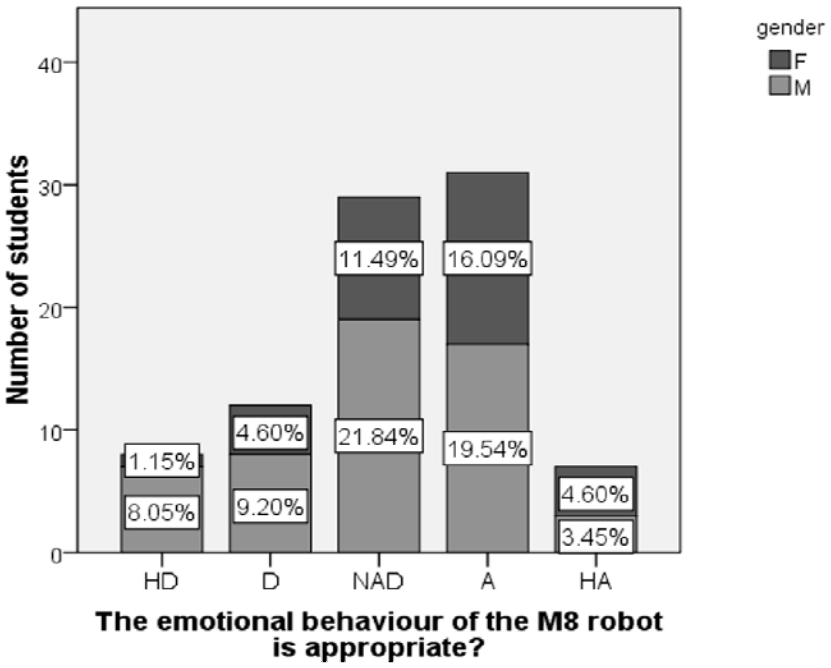


Figure 7.23 Appropriateness of M8 emotional behavior according to gender

The hypotheses are stated as follows:

H_0- There is an equal probability that the observation from one sample of students will exceed the observation from the other sample of students.

H_A- The students in the male student sample are more likely to disagree or highly disagree with the appropriateness of M8 emotional behaviour than the female student sample.

Results of the application of the Mann-Whitney U test are shown in Tables 7.41 and 7.42. Table 7.41 shows the mean rank and sum of ranks for the female and male groups tested. Table 7.42 shows the value of the statistic U and the exact significance of the one-tailed test, which is the value that is important to us, since our population is small, a exact-test should be performed and we should report the exact p-values rather than asymptotic p-values for a parametric test.

	Student gender	N	Mean Rank	Sum of Ranks
Appropriateness of M8's emotional behaviour	F	33	50.26	1658.50
	M	54	40.18	2169.50
	Total	87		

Table 7.41 Ranks corresponding to male and female student populations

	Appropriateness of M8 emotional behaviour
Mann-Whitney U	684.500
Wilcoxon W	2169.500
Z	-1.890
Asymp. Sig. (2-tailed)	.059
Exact Sig. (2-tailed)	.058
Exact Sig. (1-tailed)	.030
Point Probability	.001

Table 7.42 Statistics of the Mann-Whitney U test

In this case, the exact p-value corresponding to a one-tailed test is significant at the 0.05 level. As a result, we are obliged to accept the alternative hypothesis (H_A) and we can state that male students have a more negative perception about the appropriateness of M8-emotional behaviour in comparison to female students. Studies in Sociology and Psychology have shown that females demonstrate differences of empathy compared with males (Lennon and Eisenberg 1990), since females are expected to be concerned with harmony within the family, whilst males are expected to focus on activities that support the functioning of the so-

ciety and family. Empathy is a quality desired in a woman, since it enables her to perform her role successfully. Women are more focused on the inner space, rely more on emotional capabilities and are more intuitive than males, who are more focused on the external world.

7.5 Summary

This chapter discussed the experimental results of testing our computational emotional student model within PlayPhysics. Our model includes only observable behaviour or contextual variables to the student interaction with PlayPhysics GBL environment. The approach employs: (1) a combination of the Necessary-Path Condition algorithm with information retrieved from the Binary and Multinomial Logistic Regression and Pearson correlations to solve uncertain relations for structural learning, (2) the EM learning algorithm for parameter learning, (3) cross-validation for assessing the performance of the model on fresh data and (4) Cohen's Kappa for validating the actual agreement between student self-reports and the prospective outcome, activity and retrospective outcome emotions networks. The results of applying Cohen's Kappa for hypothesis testing showed that PlayPhysics' emotional student model is fair successful for reasoning about emotion using merely contextual random variables. However, its accuracy shows potential for refinement.

In addition, this model helped us to achieve enhanced insight into our student participants, such as observing that student confidence and attitude towards physics are significantly associated with student gender. The mouse focus, i.e. focus coarse value, and the quality of tutoring feedback perceived by students are linked to student control. The willingness to publish the outcome of a GBL environment challenge is related to the student perception of his/her performance. Also, it was observed by incorporating temporal relations, such as the past outcome, e.g. pre-test, *control t-1* and *value t-1*, enhances the capabilities of identification of PlayPhysics' emotional student model. As a result, employing a dynamic sequence of BBNs for the creation of our emotional student model not only assists us in obtaining an intelligible model, but also facilitates the representation of temporal relations that improve PlayPhysics' emotional student model performance.

When defining and assessing the activity emotions network including contextual variables and the raw value of the GSR signal, the accuracy of classification of student emotion changed from fair agreement to moderate agreement, with vaues of Cohen's Kappa of 0.358 and 0.576 respectively. The raw value of the GSR signal enhanced the recognition of student value. The raw value of the GSR signal was correlated to the independent variables: *focus coarse value, times asked help*, *value t-1* and *control t-1* after cross-validation. However, while acquiring students' GSR signal, sometimes we encountered difficulties when students have quite small and thin fingers, since the GSR sensor was not making sufficient contact, a value of 1023 was recorded that indicates open circuit. These cases were filtered from Play-

Physics database. In addition, we evaluated student experience with PlayPhysics qualitatively and quantitatively in order to gain more insight into the context and identify other potential factors, which may enhance PlayPhysics' emotional student model classification accuracy. Results of applying a t-test for independent samples showed that students in the focus and control groups enhanced their learning by interacting with PlayPhysics. Also, it was observed that students in the focus group, i.e. those who interacted with PlayPhysics GBL environment, have larger absolute and relative learning gains, i.e. with mean values of 9.620 and 0.405 respectively in comparison with the mean values of 8.000 and 0.237 from the control group. However, the difference was not significant enough to reject the null hypothesis.

From qualitatively evaluating the challenge, curiosity, storytelling and feedback elements of PlayPhysics, it was noted that approximately 51% of students 'Agree' or 'Highly Agree' that PlayPhysics' storytelling is appealing, but including other game context, such as a theme park, may encourage approximately 20% of students, who reported that they 'Disagree' or 'Highly Disagree' that PlayPhysics' storytelling is appealing. These may explain a few cases reported by students as boredom. Also, it was observed that some students found it difficult to understand the instructions given in English, since they are native speakers of Spanish, which may explain student self-reports of frustration and anger. Students mostly disagree about the appropriateness of the feedback provided by M8-robot, they think that more detailed explanations are required. In addition, investigating in more detail the difference in student perception about the emotional appropriateness of M8-robot's behaviour, it was observed that males have a more negative perception than females that may be explained by differences of empathy according to the gender.

Chapter 8: Conclusion and Future Work

Emotion has become a key research topic from the perspectives of computing and education. The fields of Human-Computer Interaction (HCI) and Affective Computing are particularly focused on identifying and reasoning about student emotion, since emotion is an important element of the human experience and because software applications that are not aware of human emotion and that cannot respond accordingly, suffer in their ability to engage users. As a result, research conducted in the area of Computer tutoring, which uses techniques such as Virtual Learning Environments (VLEs) and Game-based Learning (GBL) environments in combination with a new generation of Intelligent Tutoring Systems (ITSs) to provide interactive and personalised instruction, is focused on the ultimate goal of achieving a reasonable and reliable classification of student emotion. Simultaneously, the areas of Affective Computing and Affective Gaming are rapidly evolving. This chapter begins by providing a summary of this thesis. It proceeds to discuss our approach and findings in the light of related work. Future work and envisaged applications of the approach for the creation of other temporal data and emotional models are described.

8.1 Thesis summary

In this thesis, we presented a novel approach for emotional student modelling, which can be successfully applied in online and on-site GBL environments. We have built on ideas and strengths in related work in order to facilitate the creation of a computational model of student emotion that can achieve a fair classification accuracy of student *achievement* emotions using contextual variables and dynamic data from the previous time slice. We have shown that this accuracy becomes moderate when a combination of contextual and physiological attributes is employed. A demonstration of how a cognitive psychology theory of emotion in educational settings, Control-value theory (Pekrun et al. 2007), can be employed to create an emotional student model for reasoning about emotion in GBL environment settings was provided. We also observed that the elements of game design and instruction employed, which constitute the game-play, influence student experience of *achievement emotions*. As a result, special attention should be given to learning goals and student preferences and features related to gender.

A comprehensive review of Emotion from biological, psychological, physiological, behavioural, sociological and evolving perspectives was conducted in order to gain more insight into the factors involved in determining emotional state. The influence of emotions in education and student performance, motivation and cognitive processes was also investigated in detail. The *Control-value theory* of *achievement emotions* by Pekrun et al. (2007) was explained in detail. *Achievement emotion*s are experienced when individuals focus on performing personally meaningful activities and attaining goals, and the act of *achieving* is assessed according to predefined standards of quality. This theory states that control and value appraisals are the key to determining these emotions. *Control* is related to feeling in command, i.e. believing oneself to have and actually having the necessary knowledge and skills to be capable of performing an activity and achieving a successful outcome. *Value* is associated with the intrinsic importance of the activity or the outcome from the student viewpoint. It was noted that this theory has not previously been employed for creating a computational student model of emotion.

In addition, the field of computer tutoring was closely examined. Virtual Learning Environments (VLEs) and Game-based Learning (GBL) Environments are interactive forms of instruction that achieve personalisation when combined with Intelligent Tutoring Systems (ITSs). In both types of environments, students learn by doing and by experiencing the effects of their own actions. However, GBL environments are specifically considered more engaging and more capable of providing emotionally loaded experiences, due to providing immediate feedback through the emotional elements that constitute them, such as storytelling and sounds. The area of Affective Computing, focused on enabling machines to identify and show affective states, was introduced. The influence of Affective Computing on the areas of ITSs and GBL environments is currently encouraging research in both the identification of and demonstration of emotion. A new generation of ITSs is focused on two key problems: (1) achieving an enhanced perception of student needs, i.e. student modelling and (2) achieving adaptable instruction, i.e. tutor modelling. A detailed assessment of the strengths and weaknesses of the approaches employed by these new generation ITSs in attempting to recognise student affect and emotion was presented.

The approach that presently demonstrates the most success operates only in laboratory settings due to the requirement of expensive equipment, e.g. sensors, cameras and microphones, which are employed to acquire student data. This approach known as Cognitive-Based Affective User Modelling, which reasons about emotion from its origin, employs a cognitive psychology theory for reasoning about emotion. However, the majority of the ITSs currently employing this approach are not using a theory that has successfully explained the determination of emotion in educational settings. They have adapted existing theories of emotion to the educational context and it is not clear whether the emotional states that these

efforts attempt to identify are relevant to the teaching-learning experience. Embodied Conversational Agents (ECAs) Embodied Pedagogical Agents (EPs), Learning Companions and Affective robots are also briefly discussed as different techniques employed for conveying an emotional message. The artificial intelligence (AI) techniques employed to classify and predict student affective state and emotion by current ITSs are also described in conjunction with ways of assessing its performance.

The main hypothesis of this dissertation is that a computational and emotional student model derived and defined using Control-value theory, will achieve a reasonable accuracy of classification of relevant student emotions in GBL environments. This emotional student model comprises: a dynamic sequence of BBNs for representation of emotions and contextual variables (e.g. mouse focus and requests for help), answers to questions in-game dialogues and physiological variables, i.e. Galvanic Skin Reponse (GSR), for recognition of emotions. The formulation of our computational model was discussed and derived using Probabilistic Relational Models (PRMs). The observable random variables employed to assess emotion drawing information from the GBL environment are introduced and defined. These potential and contextual variables were identified from related research focused on determining student levels of motivation (Del Soldato and Du Boulay 1995) and self-efficacy (McQuiggan et al. 2008). The approach employed to define the structure and parameters of the network was discussed. In this investigation, the Necessary Path Condition algorithm is applied in combination with the information attained from Pearson correlations and the results of applying Binary or Multinomial Logistic Regression (BLR/MLR) in order to solve uncertain relations. The latter is employed since Bayesian models are a kind of logistic regression. Parameter learning is conducted using the Expectation Maximisation (EM) algorithm. Cross-validation is employed to determine how the defined model will perform over fresh data.

Our computational model is implemented within PlayPhysics, an emotional game-based learning (GBL) environment for teaching physics. PlayPhysics' design, implementation and functioning were described. From the analysis of lecturer and student needs, the learning and gaming objectives of PlayPhysics were set. PlayPhysics was built using Olympia architecture, which combines ITSs and GBL environments and is comprised of modules that support the provision of personalised instruction, as a reference. The implementation of Olympia's input and behaviour analysis modules were key to the acquisition of student data used to define and assess the performance of our emotional student model. The main problems addressed during the implementation of PlayPhysics were the collection of student data, e.g. contextual and physiological data, and how to facilitate and encourage student self-reporting. The Emo-report wheel was created and displayed on PlayPhysics' GUI to enable students to self-report their *achievement emotions*. The *M8-robot*, the learning companion of PlayPhys-

ics, serves the purpose of asking for student self-reports in addition to providing hints about student misconceptions when appropriate. A version of PlayPhysics, as an on-site GBL environment, was created in order to support the recording of the student Galvanic Skin Response (GSR) signal using a Bluetooth sensor, which was created using the LEGO NXT intelligent brick and LEJOS. PlayPhysics was implemented using principally Java, the Unity 3D game engine and JavaScript.

The results of applying the proposed approach for defining and evaluating PlayPhysics' emotional student model using only contextual variables and a combination of contextual and physiological random variables were presented. The approach for facilitating the derivation and evaluation of the PlayPhysics' emotional student model using a dynamic sequence of BBNs gave promising results. We do not present the agreement between student self-reports and PlayPhysics' emotional student model predictions using only percentages, since we believe that approach can be misleading. For this reason, we employ Cohen's Kappa to validate our alternative hypothesis. Results were shown to be promising and not to be random. It was observed that using the Control-value theory to reason about student *achievement emotions* achieves a *fair* accuracy of classification using only contextual variables, covering a range of values of Cohen's Kappa that is larger than 0.2 and lower than or equal to 0.4. However, when physiological variables, e.g. GSR signal, are employed in combination with contextual variables a *moderate* accuracy of classification is attained, covering a range of values of Cohen's Kappa larger than 0.4 and lower than or equal to 0.6.

The resulting emotional student model has the potential for enhancement, since the most accurate diagnosis systems achieve a value of Cohen's Kappa of at least 0.75. However, the resultant emotional student model is intelligible, provides more insight into factors actually related to *control* and *value* and, as a result, helps in determining student emotion. For example, it was noted that student confidence is related to gender, i.e. the majority of female students show lower levels of confidence related to achieving a successful outcome when compared with male students. Also, the raw value of the GSR signal demonstrated a medium strength association with mouse focus and the number of times a student asks for help, i.e. cognitive states of concentration and clarification. It was also sometimes related to the *control t-1* and *value t-1*, dynamic variables from the past time slice. In addition, the precision and specificity of classification for each category of *achievement emotions*, *control* and *value* was reported.

Furthermore, in order to gain more insight into the features of the GBL environment that influence student experience of *achievement emotions,* a quantitative and qualitative evaluation of PlayPhysics was conducted. Results showed that students interacting with PlayPhysics achieved positive absolute and relative learning gains. Specifically students belonging to the focus group, who interacted with the PlayPhysics' GBL environment, achieved larger

mean values corresponding to absolute and relative learning gains (9.620 and 0.405 respectively), but not significantly larger than gains obtained by students in the control group, who used a PowerPoint presentation in a video format in PlayPhysics to achieve the same knowledge about the relevant physics concepts. It was noted that the majority of students agreed and highly agreed with the elements of challenge, curiosity and usability of the PlayPhysics GBL environment. Also, it was observed that the majority of male students disagree with *M8-robot* emotional behaviour in comparison with the majority of female students, who have a tendency for agreeing with it. This may be due to known differences of empathy associated with gender.

8.2 Relation to other work

This section discusses our results and approach applied in relation to other work. PrimeClimb (Conati and Maclaren 2009) (see Chapter 3,Table 3.2) relates closely to PlayPhysics, which also combines GBL environments and a new generation of ITSs and is centred on identifying student emotion. However, PrimeClimb was originally created to teach Mathematics to children at the primary school level, but was evaluated with undergraduate students, whilst PlayPhysics is created to teach physics to students taking an introductory course of Physics, students in their last year of High school and their first years of undergraduate level and is evaluated with this target population. It is important to note that the intensity and experience of emotion changes with student age (Alea et al. 2004).

The approach employed is a combination of identifying the physical/physiological effects of emotion and reasoning about emotion from its origin. In our case, we focused mainly on assessing the approach of reasoning about student emotion from its origin and we compared the performance of this approach with the Hybrid approach employed by Conati and Maclaren (2009). It was observed that the accuracy of the model increased by incorporating physiological signals. The cognitive psychological theory employed by this group to derive their DBN was the Ortony, Clore and Collins (OCC) model (Ortony et al. 1990). This theory was not originally created to explain emotion in an educational context, but applied for reasoning about emotion from personal diaries. As a result, even though the theory was adapted to the learning context, it is not clear whether the emotions chosen are relevant to the learning experience. Conati and Maclaren (2009) employed an EPA to remind students to self-report their emotional state. PlayPhysics employs the learning companion *M8-robot* for this purpose. Students interacting with PrimeClimb can use a pop-up window to report their emotion. PlayPhysics employs the Emo-report wheel for this purpose. Joy, distress, admiration and reproach are the emotions identified by PrimeClimb. Conati and Maclaren (2009), as well as other research groups, presented their results as percentages of agreement between student self-reports and the predictions of their emotional student models (69.59%, 62.30%,

67.42% and 38.66% accuracy for joy, distress, admiration and reproach respectively), which makes it difficult to quantify its reliability.

The CRYSTAL ISLAND learning environment (Sabourin et al. 2011) also focuses on recognising achievement emotions as does PlayPhysics. However, it employs the appraisal-based theory of learning emotions by Elliot and Pekrun (2007). This theory is different from Control-value theory in that it links the attainment of performance or mastery of goals and its valence with the experience of achievement emotions in a learning context. Similar to our investigation, Sabourin et al. (2011) do not include the category *no-emotion* in their model, since it is not defined by this theory, similarly, this is not defined in Control-value theory. The results reported are also in the form of percentages of agreement, so it cannot be assessed whether the agreement is random or otherwise. Sabourin et al. (2011) assess different AI techniques in order to find the most suitable representation of their model. In their research DBNs were the technique that achieved the best performance. However, they focused on identifying students' confusion, curiosity, excitement, focus, anxiety, boredom and frustration. The latter two were identified with accuracies of 18% and 28% respectively, whilst PlayPhysics identifies these emotions using Control-value theory with accuracies of 20% and 60% respectively using only contextual variables. Our emotional student model is the first and only model to date that was implemented using Control-value theory.

Another theory that is not related to the learning context, but has been adapted and employed to identify emotions in the educational settings using facial expressions is the theory by Ekman (1999), which has shown to be quite effective when employed by Autotutor (D'Mello et al. 2008b) in laboratory settings. However, the approach has still not demonstrated its effectiveness in the place where learning commonly happens, the classroom or via online. Autotutor uses AI techniques, such as artificial neural networks, to classify the effects/features of emotion, which do not result in an intelligible computational model of emotion. These models are a kind of black box where inputs and outputs are known, but do not provide enhanced information about their relationship, the participants or the affective context. PlayPhysics' emotional student model is an intelligible model that assists us in identifying factors that are considered actual predictors of *control* and *value* and the manner in which these are associated. The model also assists us in achieving an enhanced understanding of the student population.

The application of Probabilistic Relational Models (PRMs) is employed in this dissertation to achieve an enhanced understanding of the random variables of the domain that may be considered for the creation of our emotional student model. As a result, they facilitate the definition of Bayesian student models. This approach has been employed previously by Sucar and Noguez (2008) for understanding the domain involved in defining a student model capable of identifying a student's level of knowledge and understanding. We are the first to

employ this approach in modelling an affective domain. The application of the Necessary Path Condition algorithm for structural learning has been successfully employed in the area of telecommunications (Bashar et al. 2010) when scarce data is available. In this dissertation this approach is employed in combination with information acquired through applying Binary and Multinomial Logistic Regression and Pearson correlations in order to solve uncertain relations. Pearson Correlations have been successfully employed previously as a criteria for creating the structure of a Bayesian student model of student attitudes (Arroyo and Woolf 2005). Results of applying Binary and Multinomial Logistic Regression are employed as a criteria or guidance for creating the network structure, since Bayesian models are a kind of Logistic Regression (Roos et al. 2005) and we can determine the contribution of each selected variable to the prediction. We are not aware of any other research that employs Binary or Multinomial Logistic Regression for the purpose employed in this dissertation.

The contextual and observable behaviour variables employed in this research have been employed previously by research that endeavours to identify student levels of motivation (Del Soldato 1993) and self-efficacy (McQuiggan et al. 2008), which were identified as being related to the factors employed by Pekrun et al. (2007) to assess control and value. Hence, these variables were identified as potential predictors and employed in this dissertation to recognise student *achievement emotions*. The design of a GSR sensor built using the LEGO intelligent brick by Gasperi (2010) was adapted to allow connections via Bluetooth. In this thesis this sensor was effectively employed in on-site GBL environments for acquiring real time data corresponding to student physiological signals.

8.3 Future work

There are many potential areas of future work for our computational model and PlayPhysics implementation. Our emotional student model can be applied more generally to any kind of learner, i.e.not necessarily a student. Also, the model achieves a *fair* accuracy of classification when using only contextual variables for recognition and improved accuracy, i.e. *moderate*, when employing a combination of contextual and physiological variables for recognition. However, our emotional student model still does not reach the accuracy of diagnosing systems considered highly reliable, i.e. $\kappa = 0.75$. It is still necessary to identify other factors or independent variables that can enhance the accuracy of classification of *control* and *value* for certain categories. Hence, future work will investigate features in the GSR signal that can be employed to provide an enhanced accuracy of classification, such as the average of the GSR value (McQuiggan et al. 2008) or its kurtosis and skewness. The latter has proven useful in deciding between positive and negative emotions (Rajae-Joordens 2008). Also, facial expressions, employing computer vision, can be incorporated to enhance the accuracy of classification of value, since this information has proven quite affective for this purpose (Arroyo et

al. 2009). Natural language processing to Identify student attitude, i.e. sentiment analysis, and speech intonation features could also be explored (D'Mello et al. 2008a). In addition, future work can also investigate the manner of identifying the base-line or Neutral emotion, which is a limitation of Control-value theory.

Whilst PlayPhysics' emotional student model performs with reasonable accuracy, it was noted in Chapter 7 that the results of control and value appraisals are fed back (*control t-1* and *value t-1*) instead of the resultant achievement emotion, i.e. *emotion t-1*). This deviates from the original theory defined by Pekrun et al. (Pekrun et al. 2007) and may constitue a limitation of our emotional model. Therefore, future work may be focused on assessing the accuracy of our model by feeding back the causal effect of the inferred achievement emotion.

PlayPhysics itself can be enhanced with more detailed feedback to students and *M8-robot* can also provide more instruction together with affective feedback. In addition, other game challenges or fantasies can be included in order to appeal more to students who perceive the space adventure theme as unattractive. Furthermore, the Bluetooth GSR sensor employed could be improved, since sometimes the hoop and loop bands including the cupper electrode, which are connected to the student's index and middle fingers of the less favoured hand, are less effective when students have fingers that are very thin and narrow. Other methods of acquiring the GSR measurement could be explored. For example, a GSR sensor that is becoming increasingly popular is the wearable Q sensor created by *Affectiva (2012)*, a company spun out from the Massachusetts Institute of Technology (MIT). This Q sensor is a comfortable wrist band that captures user physiological data.

The methodology for derivation of our dynamic emotional student model can be applied to deriving other dynamic emotional models that can be employed in application areas other than education. For example, in an e-Commerce context, an emotional model could identify relevant emotional states of consumers in order to select suitable marketing techniques for selling products online (Defino 2011). Our approach can create dynamic data models in other fields, such as Biology, where the representation of dynamic models of cells are key to achieving an enhanced understanding of Genetics (Falcke 2012).

8.4 Conclusion

The key purpose of this research was to investigate whether the creation of a computational model of student emotions employing Control-value theory (Pekrun et al. 2007), can obtain a reasonable accuracy for recognition of student emotions in online and on-site GBL environment settings. PlayPhysics was the implementation employed to test whether our emotional student model can be applied to GBL environment settings. Results showed that the model achieves a *fair-moderate* accuracy of classification using only contextual variables and a *moderate-substantial* accuracy using a combination of contextual and physiological vari-

ables. However, the resultant model is not as reliable as highly accurate diagnosis systems $(\kappa = 0.75)$. Hence, future work will focus on finding other random variables, such as other GSR response features and analysis of facial expressions and speech intonation features, which can assist in classifying *control* and *value* more accurately. The approach employed to facilitate the derivation of a dynamic sequence of BBNs proved effective in creating an intelligible emotional student model. This approach can be employed to derive other dynamic and intelligible data models that can assist in achieving enhanced understanding in other contexts and research areas other than education, such as e-Commerce and Genetics. Therefore, it is envisaged that this research can have a significantly impact in research areas, which share the goals of adapting to the user and gaining insight from a domain through the creation of data models, in addition to the prospective areas of Affective Computing, Student Modelling and Computer Tutoring.

Appendices

Appendix A. Online survey for requirements gathering

Here are presented the two online surveys that were made for gathering lecturers and students' needs at ITESM-CCM and Trinity College Dublin. The survey gathering students' data is shown as follows:

A.1 Student questionnaire

Survey
Motivating the learning of Physics

Student's Questionnaire

Thank you for responding to this survey. Your answers will help us to develop enhanced computer technologies that can assist your understanding and encourage your motivation for studying physics.

1. Sex

☐ Female ☐ Male

2. Please, select your age range

☐ Less than 18 ☐ 18 – 20 ☐ 21 - 23 ☐ 24 - 25 ☐ More than 25

3. Please, select your university

☐ Trinity College Dublin ☐ Queen's University Belfast ☐ Tecnológico de Monterrey-CCM

4. What is the title of the undergraduate course that you are studying?

Personal Preferences for learning

5. Please order the learning styles that you prefer from 1 to 4, where 1 is "very comfortable", 2 is "comfortable", 3 is "fairly comfortable" and 4 is "not comfortable".

☐ Auditory (I prefer to learn from oral presentations, storytelling or speech)

☐ Visual/ Non verbal (I prefer to learn from pictures, charts, etc.)

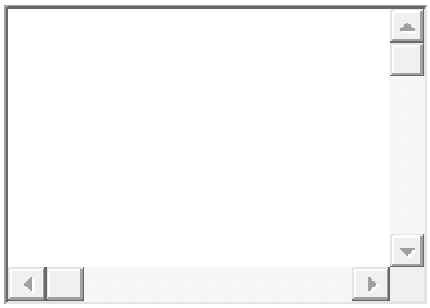
☐ Tactile/ Kinaesthetic (I prefer to learn by doing)

☐ Verbal/Visual (I prefer to learn from written words)

6. Please compare the statements on left and right and select the value on the scale 1-10, which corresponds to the position that best describes where your preference lies.

I like working in a team	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10	I like working largely alone
I like days full of action	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10	I like days when I have a chance to reflect, write or study
I like practical work, requiring precision, and using hard data	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10	I like using data to draw out meanings and develop theories and possibilities
I like to see short term results from my work	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10	I am happy with delayed results if I can see a long term vision
I am best at making decisions using logical, objective analysis	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10	I am best at making decisions based on what is important to individuals
I like to be complimented on my competence	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10	I like to be complimented on my qualities as a person
I like creating organisation and structure	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10	I like being flexible and spontaneous
I like to know what I am doing and when	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10	I like spontaneity and the feeling of unexpectedness
I prefer it when I have made a decision	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10	I like to keep my options open

7. From your perspective, what are the three main topics of your Physics module that are considered to be difficult and why?



Computer equipment and skills

8. Please, select the statements that describe most accurately the University's computer equipment.

- ☐ Responds quickly to user events (It opens quickly an internet window)
- ☐ Responds moderately to user events
- ☐ Responds slowly to user events
- ☐ It has microphone capability
- ☐ It has speakers or earphones
- ☐ It has web cameras

9. Do you have a computer at home?

☐ Yes ☐ No

9.5. If yes, can you please select the statements that best describe this equipment?

- ☐ Desktop
- ☐ Laptop
- ☐ Responds quickly to user events (It opens quickly an internet window)
- ☐ Responds moderately to user events
- ☐ Responds slowly to user events
- ☐ It has microphone capability
- ☐ It has speakers or earphones
- ☐ It has web cameras

10. Approximately, what age were you when you started to use computers?

☐ age

11. While using a computer in general, I feel…

- ☐ Very confident
- ☐ Confident
- ☐ Neutral
- ☐ Not confident
- ☐ Extra not confident

12. While using the Internet I feel…

- ☐ Very confident
- ☐ Confident
- ☐ Neutral
- ☐ Not confident
- ☐ Extra not unconfident

13. Please, select the expression that best describes your experience at playing computer games

- ☐ I have never played a computer game
- ☐ I hardly ever play computer games
- ☐ I occasionally play computer games
- ☐ I usually play computer games
- ☐ I always play computer games

Education

14. Please order the features of an educational computer game given below according to their influence to allow you to feel engaged and to facilitate learning (1 is the most important and 8 is the least).

- ☐ Player Roles
- ☐ Game Rules
- ☐ Goals and objectives
- ☐ Puzzles, problems or challenges
- ☐ Narrative or Story
- ☐ Player's interaction
- ☐ Payoffs and strategies
- ☐ Outcomes and feedback

15. Please give an example of personal experiences involving one or more of the elements given in question 14 that helped you to feel engaged

16. From your perspective, please order the elements given below according to their influence for feeling engaged (1 is the most relevant and 8 is the least).

- ☐ Colours
- ☐ Graphics
- ☐ Sounds
- ☐ Emotional responses or behaviour
- ☐ Empathy
- ☐ Believability
- ☐ Entertainment or sense of humour
- ☐ Provision of challenges

17. From your perspective, please order the elements given below according to their influence for learning concepts (1 is the most relevant and 8 is the least).

- ☐ Increase the level of difficulty
- ☐ Introduction of new strategies
- ☐ Payoffs
- ☐ Embody concepts to be learned
- ☐ Enquiries
- ☐ Explanations
- ☐ Clues
- ☐ Interactive dialogues with animated characters

18. Please associate each colour(letter) with an emotion without repeating letters.

A. Blue	1. ☐ Neutral (Without emotion)
B. Pink	2. ☐ Happy

C. Purple
D. Green
E. Yellow
F. Black
G. Red
H. Orange
I. Khaki
J. Gray
K. Brown

3. ☐ Sad
4. ☐ Disgust
5. ☐ Fear
6. ☐ Surprise
7. ☐ Anger

19. If you were asked to pick a colour to describe frustration, which colour would it be?

☐

Return

A.2 Lecturer questionnaire

The survey enquiring lecturers' viewpoint is shown as follows:

Lecturer's Questionnaire

Thank you for responding to this survey. Your answers will help us to develop learning environments that assist the learning and teaching of physics. In addition, they can help to measure the student's performance and provide personalized feedback.

1. Please, select your university

☐ Trinity College Dublin ☐ Queen's University ☐ Tecnológico de Monterrey-CCM

2. Please state the title of your module

3. Approximately, how many students are there in your Physics module?

0

4. Approximately, what percentage of students fail the module each semester or term?

00.00 %

5. What are the three main reasons for failing?

6. Which topics of the module are considered to be difficult and why?

7. Which topics of the module are considered to be fundamental and why?

8. To deliver the module, what are the current and most commonly used methods and resources? (e.g. tutorials, readings, exams, etc.)

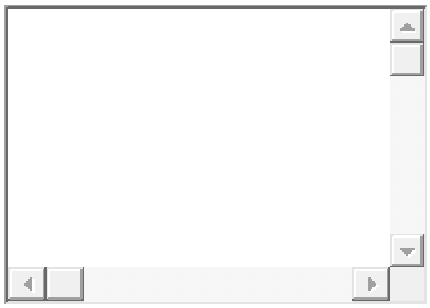
9. Do you consider these methods and resources sufficient to attain the main goals, objectives, knowledge, skills and abilities required by the module?

☐ Yes ☐ No

10. Do you consider that most students (80%) can attain the main goals, objectives, skills and abilities of the module in the available time?

☐ Yes ☐ No

11. From your perspective, if you were able to design a computer game that allows your students to learn the most difficult topics of your module, what ideas come to mind? What kind of story, game genre, player roles, game rules would you suggest?



Return

A.3 Results for students at ITESM-CCM

Here are presented additional graphs and tables corresponding to the analysis of students' needs and preferences for learning at ITESM-CCM.

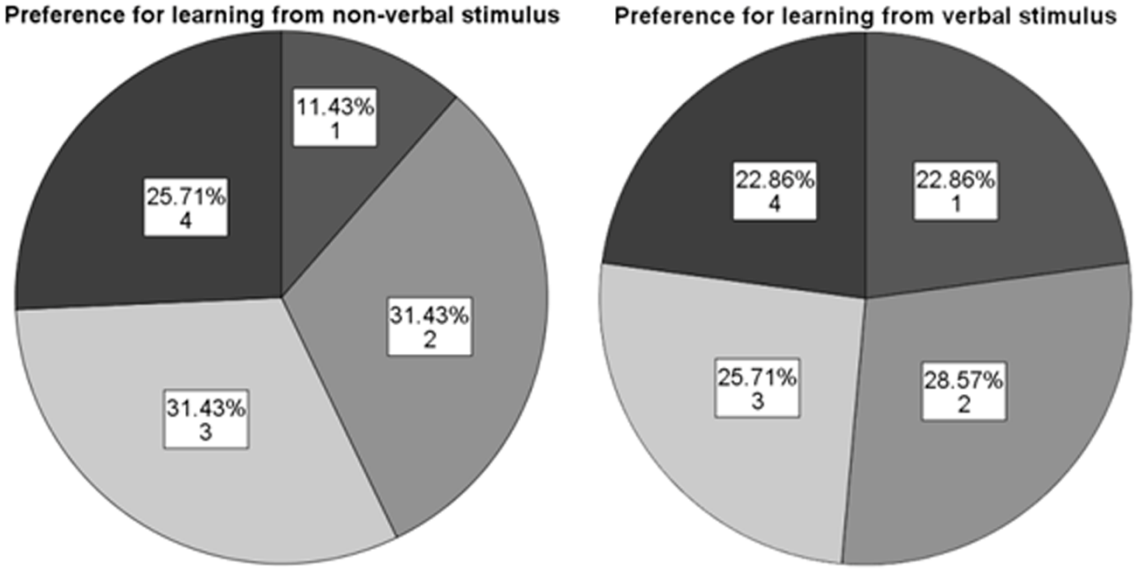


Figure A.1 ITESM students' preferences for learning from stimuli

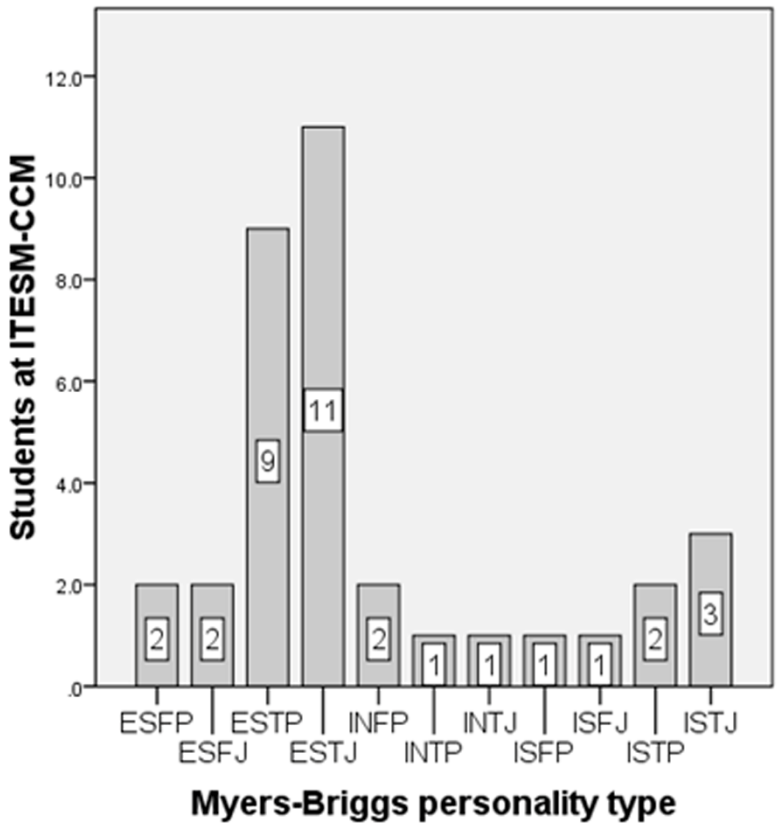


Figure A.2 ITESM students' Myers-Briggs personality types

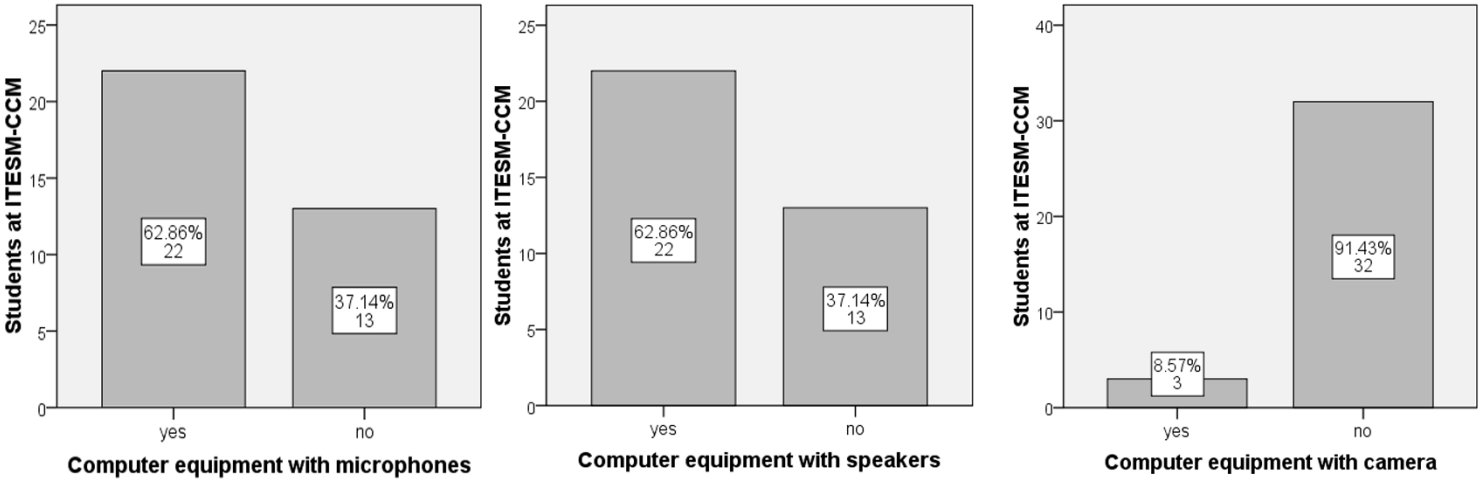


Figure A.3 Graphs describing the available computer equipment at ITESM-CCM

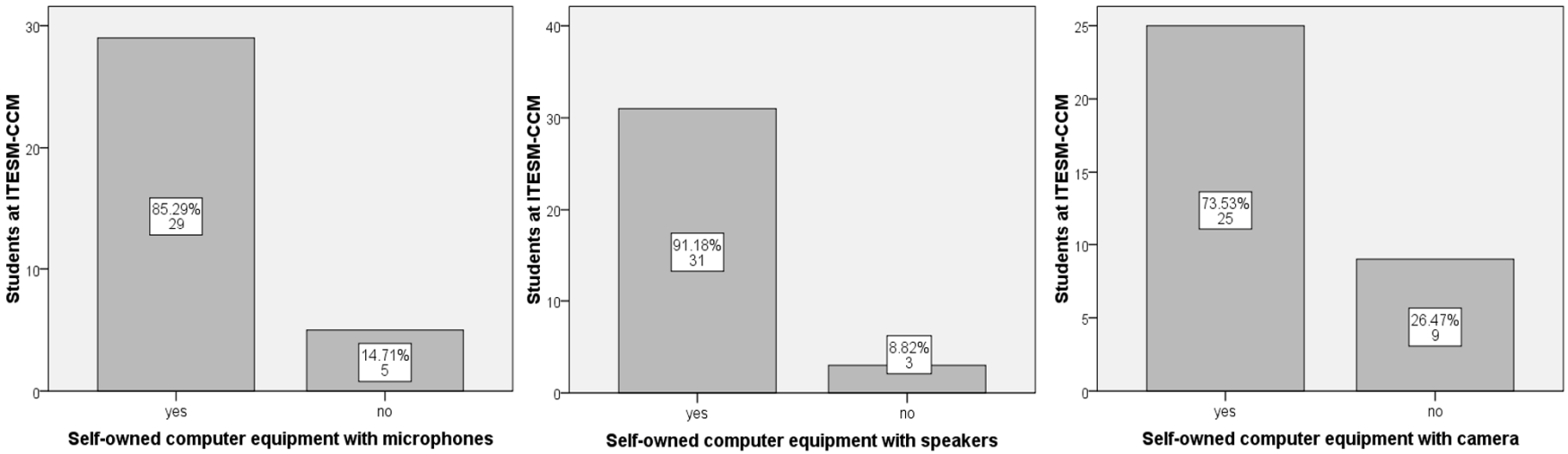
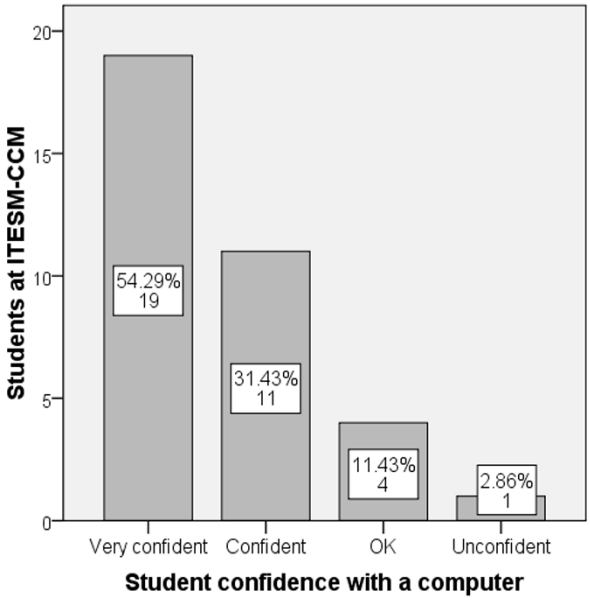


Figure A.4 ITESM-CCM students self-owned computer equipment features



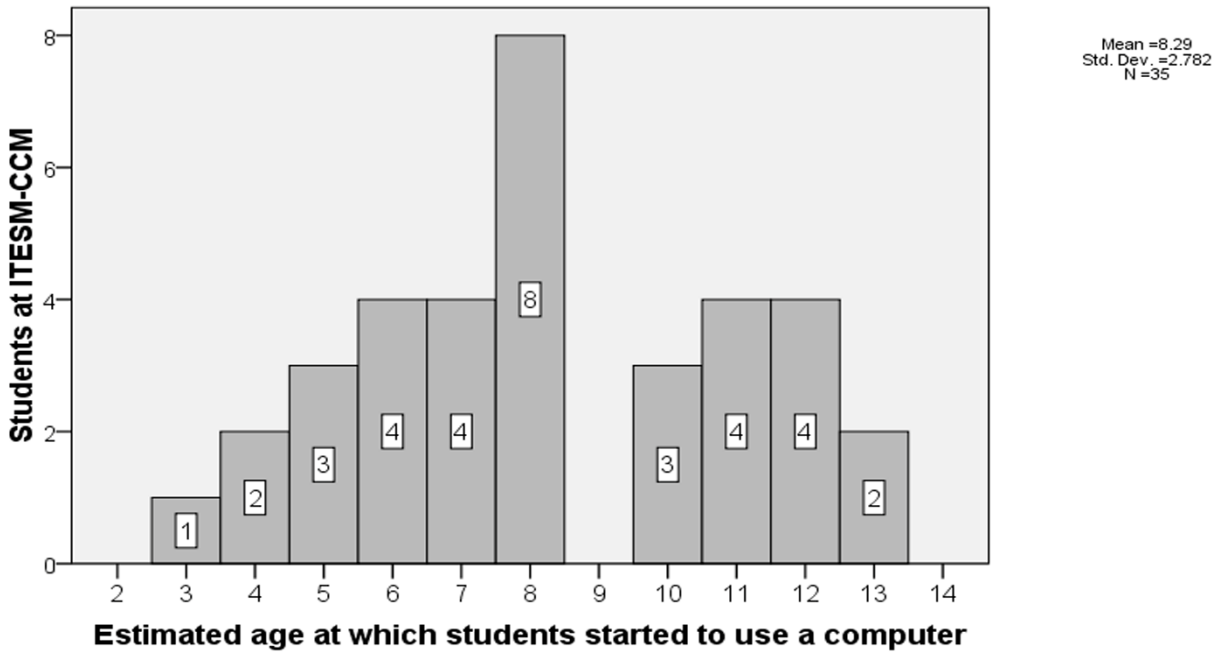


Figure A.5 Graphs illustrating student confidence and familiarity with computers

		Player roles	*Game rules*	*Goals and objectives*	*Puzzles, problems or challenges*	*Narrative or Story*	*Player's interaction*	*Payoffs and strategies*	*Outcomes and feedback*
N	*Valid*	35	35	35	35	35	35	35	35
	Missing	0	0	0	0	0	0	0	0
Mean		4.660	5.000	3.740	3.860	4.690	3.770	4.570	5.710
Std. Deviation		1.748	2.623	2.020	2.002	2.978	2.088	2.062	2.080

Table A.1 ITESM-CCM student perceived importance of game features

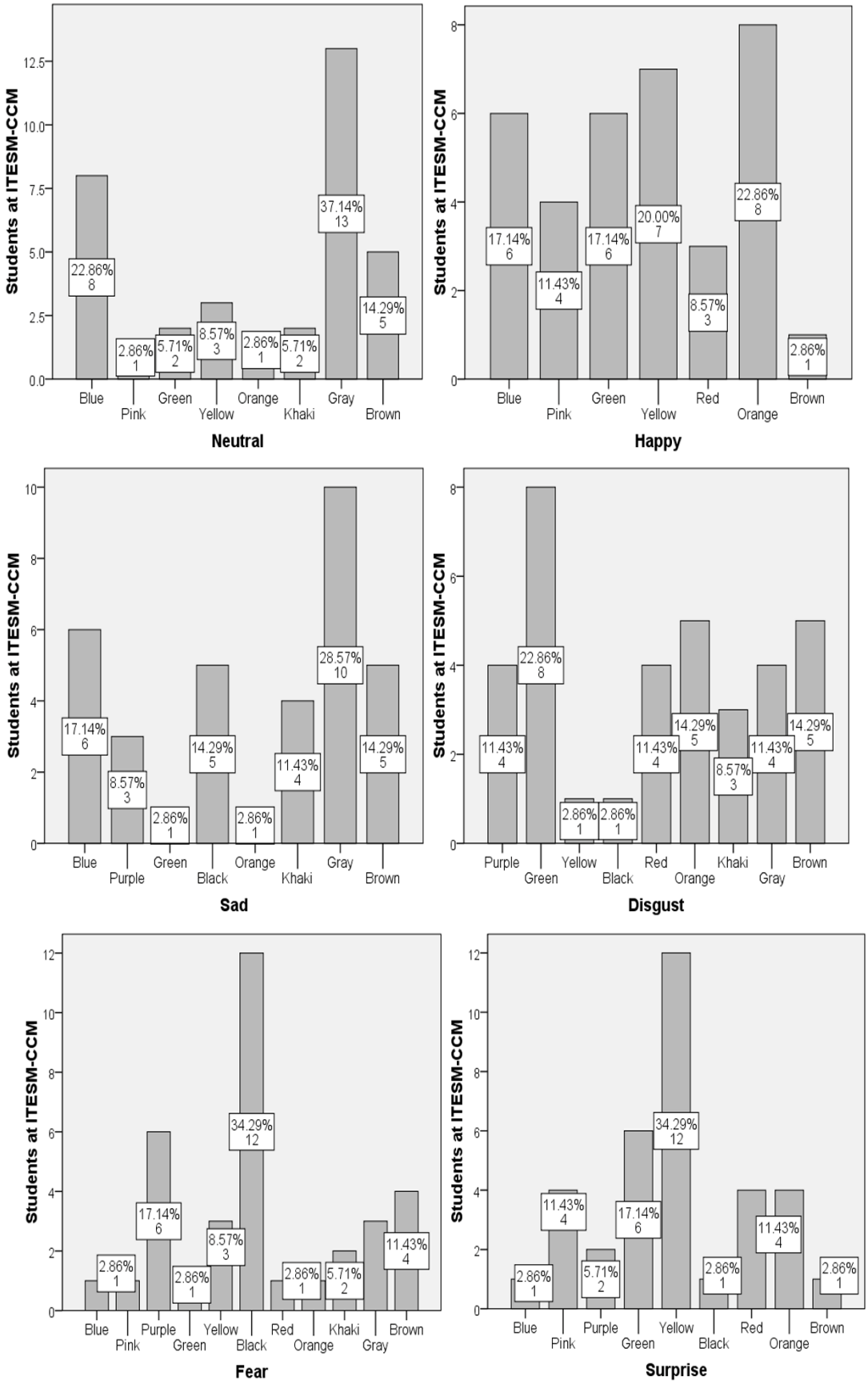


Figure A.6 ITESM-CCM student associations between colours and emotions

A.4 Results for students at Trinity College Dublin (TCD)

Here are presented additional graphs and tables corresponding to the analysis of students' needs and preferences for learning at TCD.

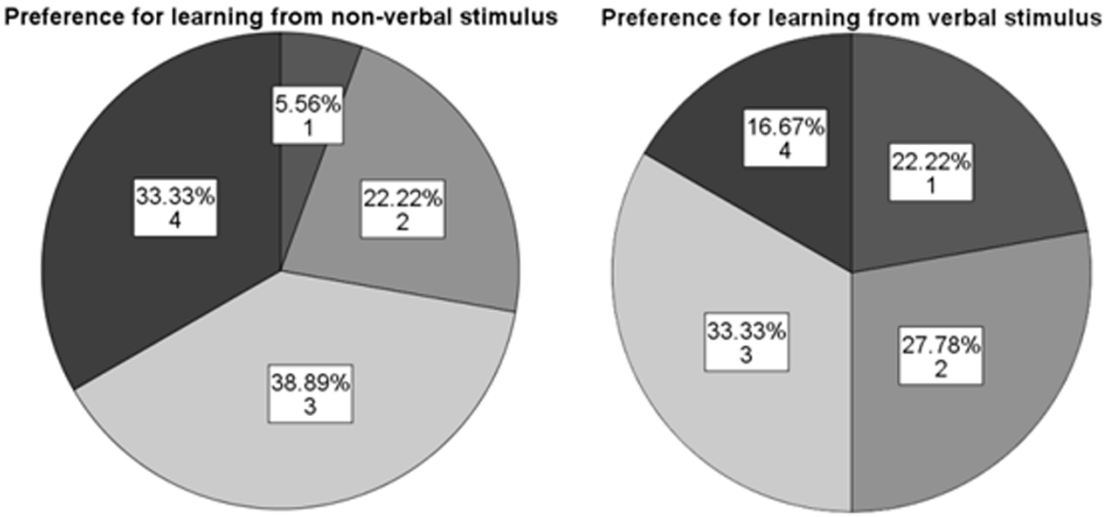


Figure A.7 TCD students' preferences for learning from stimuli

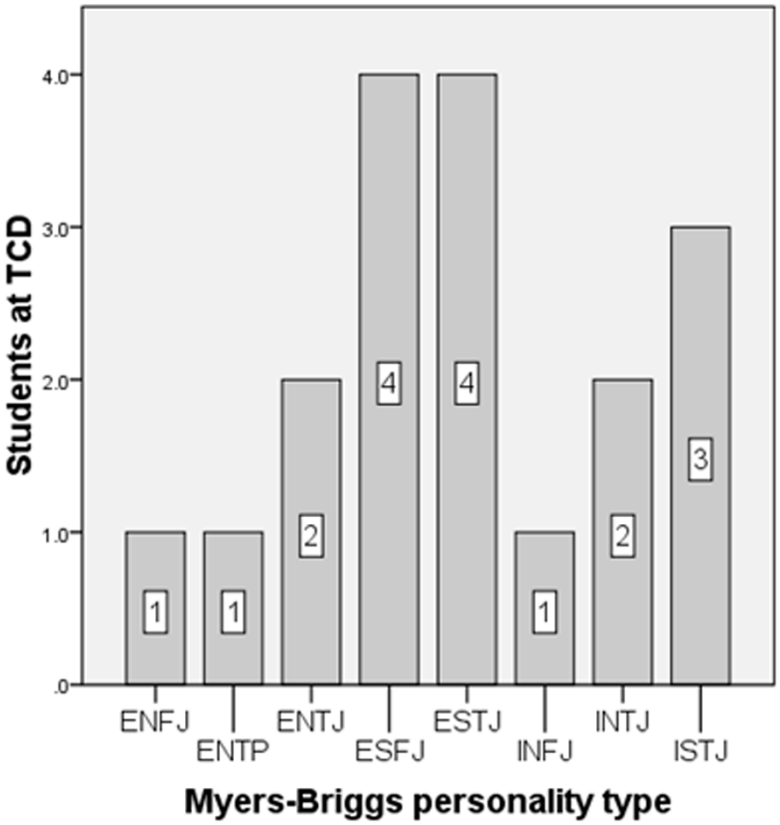


Figure A.8 TCD students' Myers-Briggs personality types

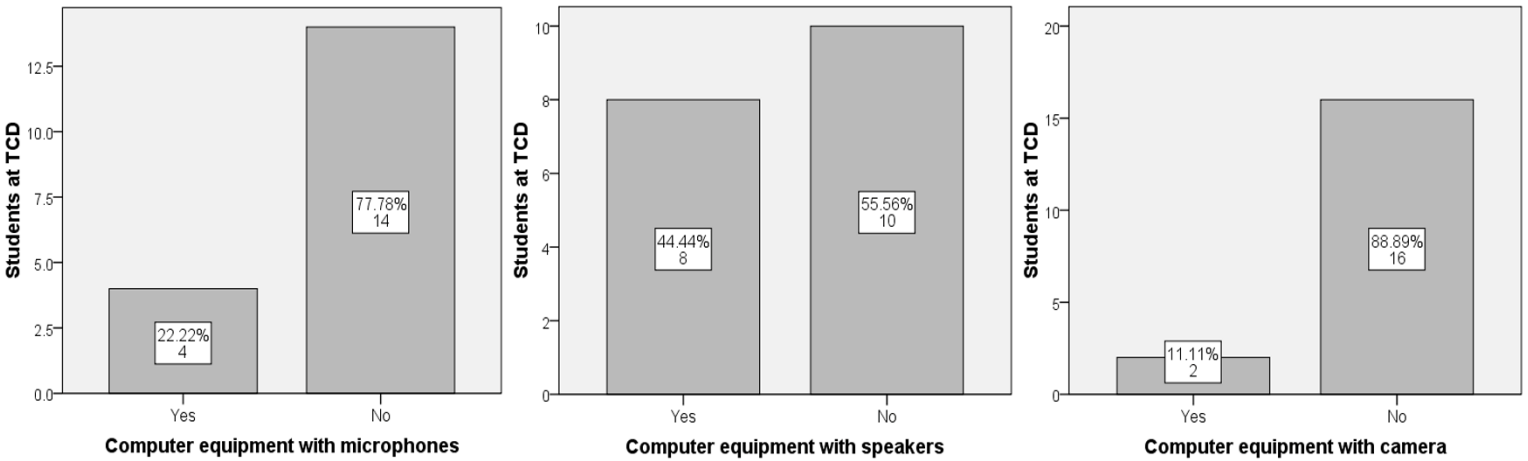


Figure A.9 Graphs describing the available computer equipment at TCD

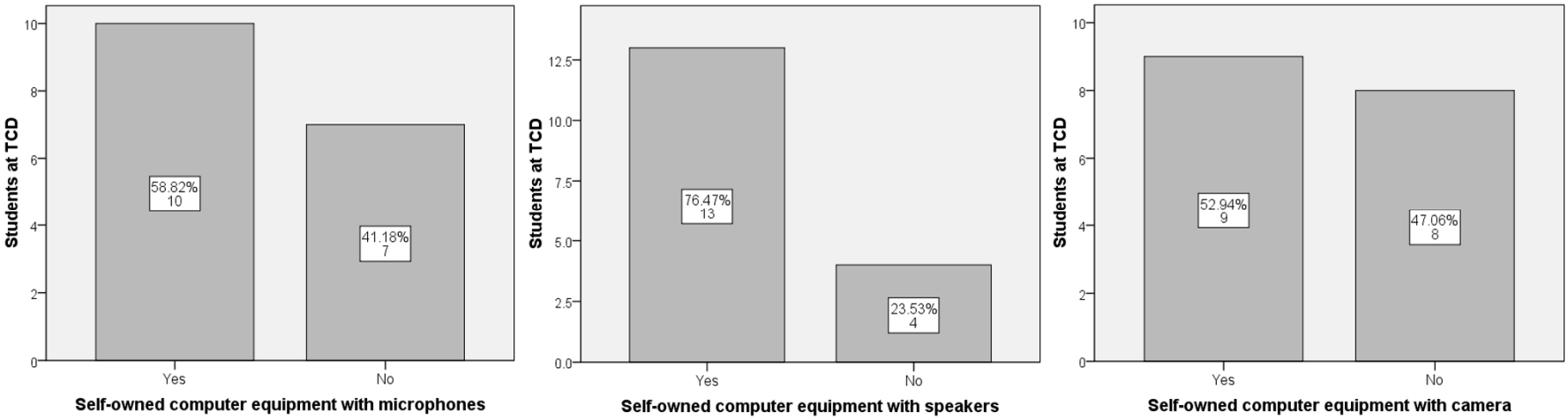
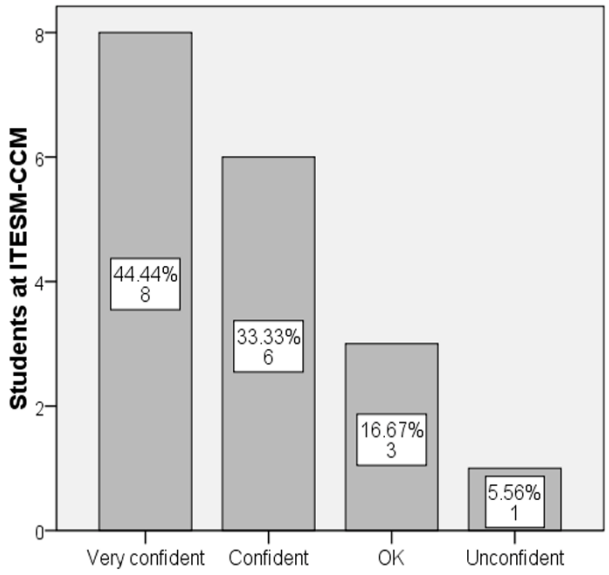


Figure A.10 TCD students self-owned equipment features



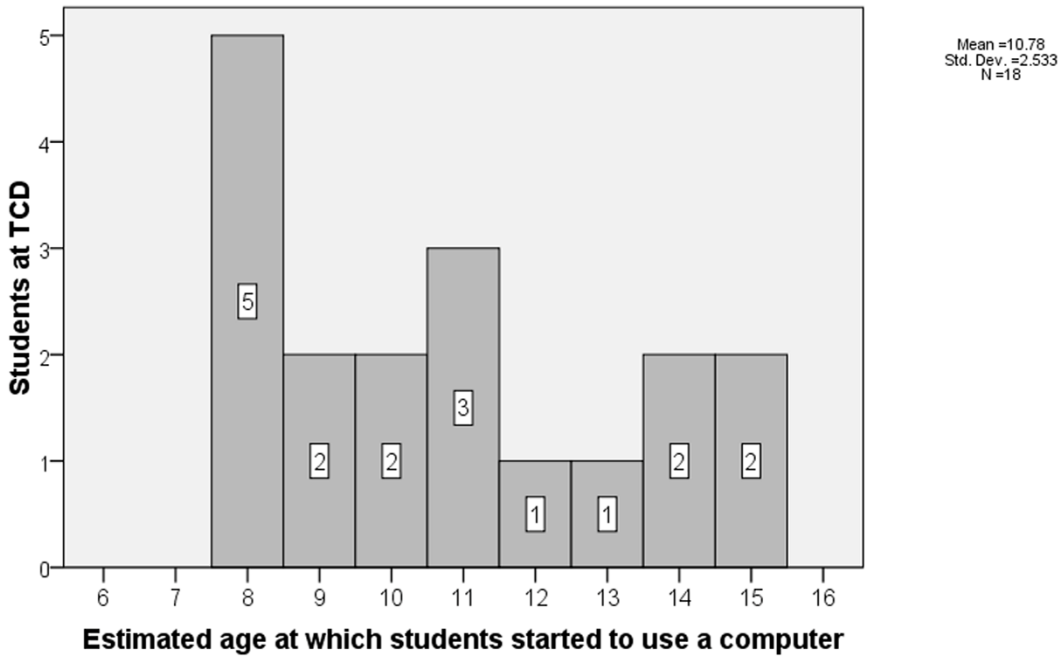


Figure A.11 Graphs illustrating student confidence and familiarity with computers

	Player roles	*Game rules*	*Goals and objectives*	*Puzzles, problems or challenges*	*Narrative or Story*	*Player's interaction*	*Payoffs and strategies*	*Outcomes and feedback*
N Valid	18	18	18	18	18	18	18	18
Missing	0	0	0	0	0	0	0	0
Mean	3.830	5.940	3.780	3.000	4.830	3.940	5.110	5.560
Std. Deviation	2.728	1.893	1.665	2.000	2.503	2.014	1.605	2.479

Table A.2 TCD student perceived importance of game features

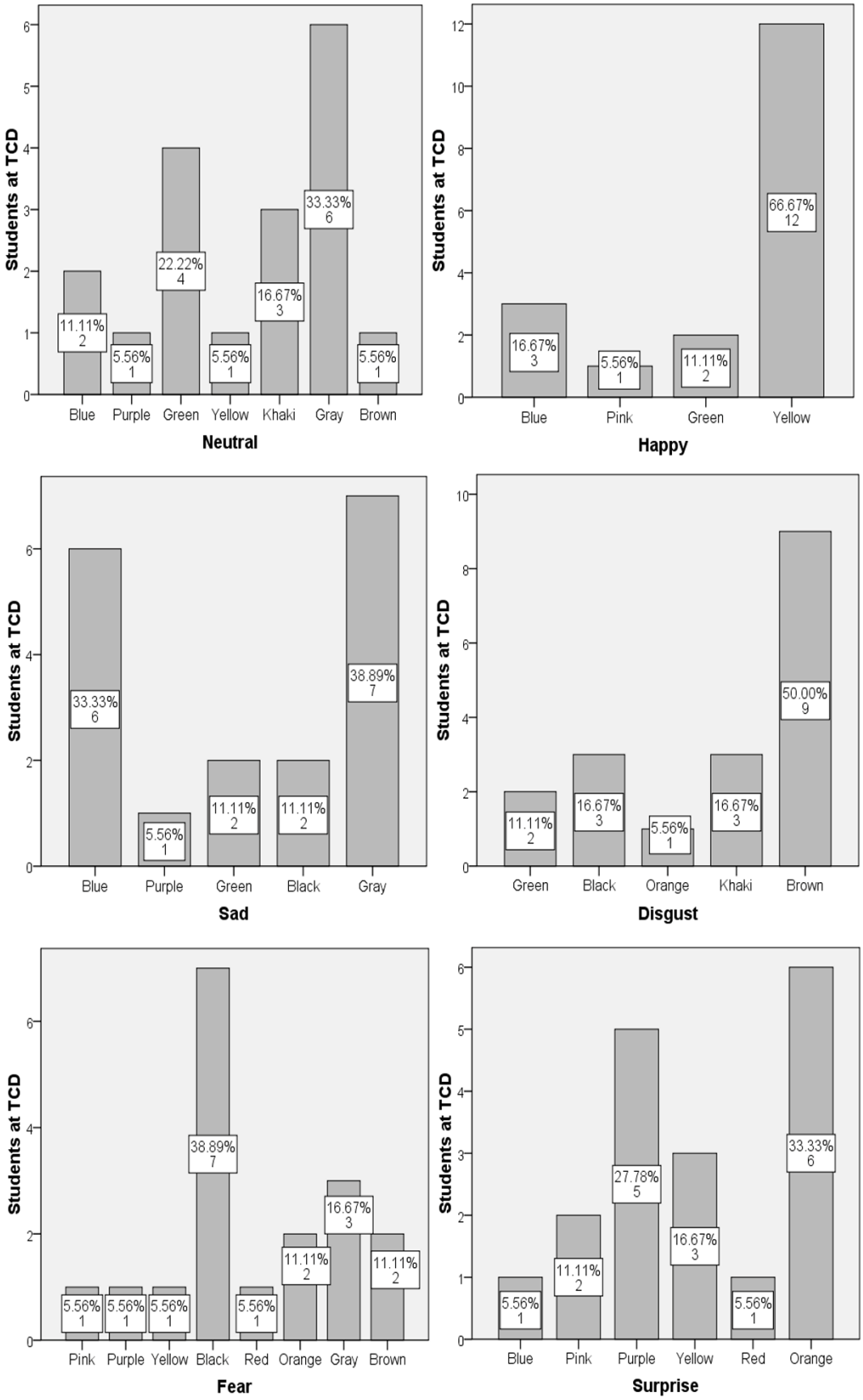


Figure A.12 TCD student associations between colours and emotions

Appendix B. PlayPhysics storyboard

The lecturers at ITESM-CCM suggested reading the book by Serway and Jewett (2004), used as the main reference for delivering the course of Physics I. Chapters 1 to 12 were read and examined in order to derive some ideas for defining PlayPhysics game challenges. As a result the storyboard shown in Figure B.1 was created.

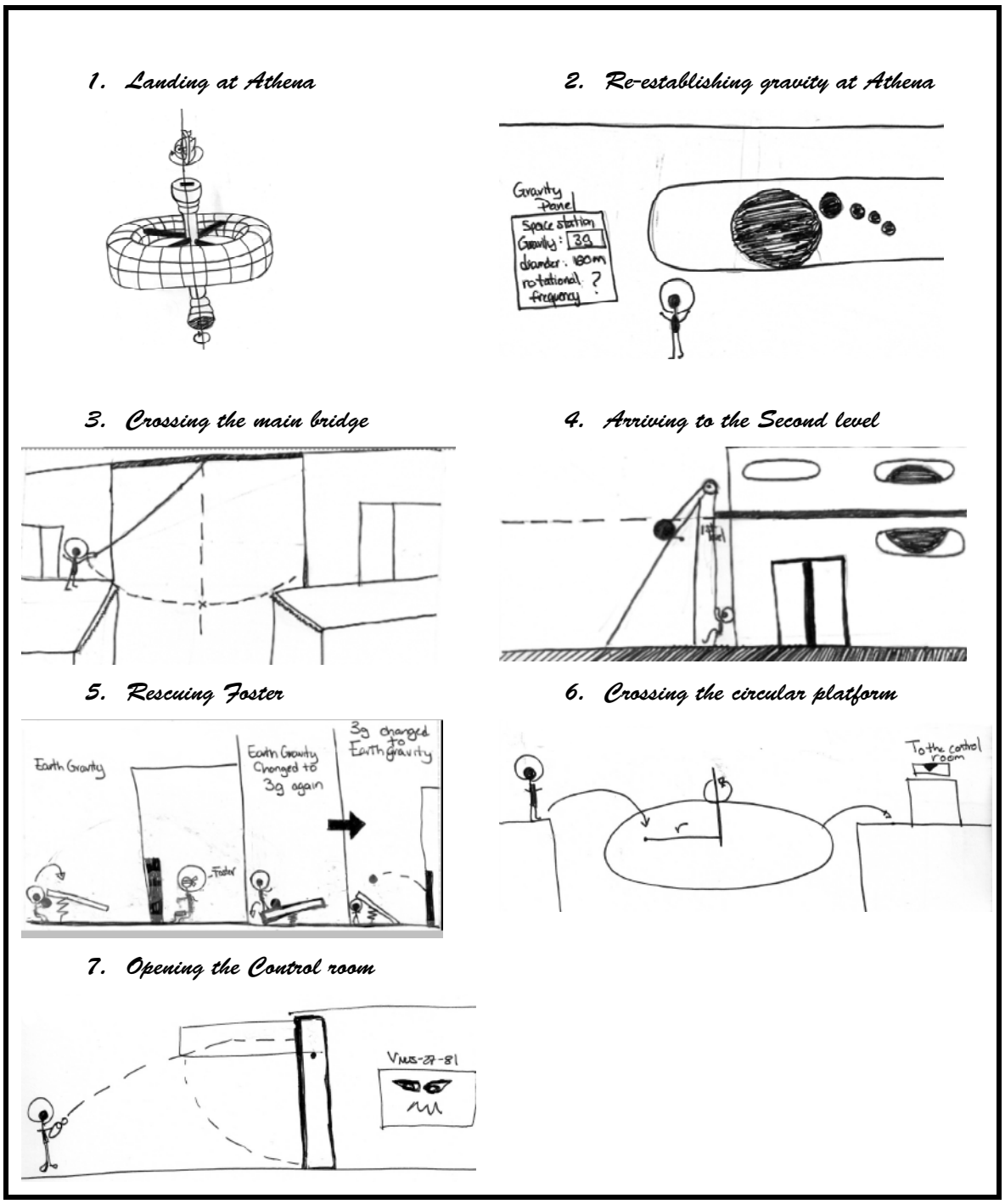


Figure B.1 Storyboard of the proposed game challenges designed for PlayPhysics

Appendix C. PlayPhysics player characters

This section shows additional material, which was created in order to support the design and implementation of PlayPhysics' player characters.

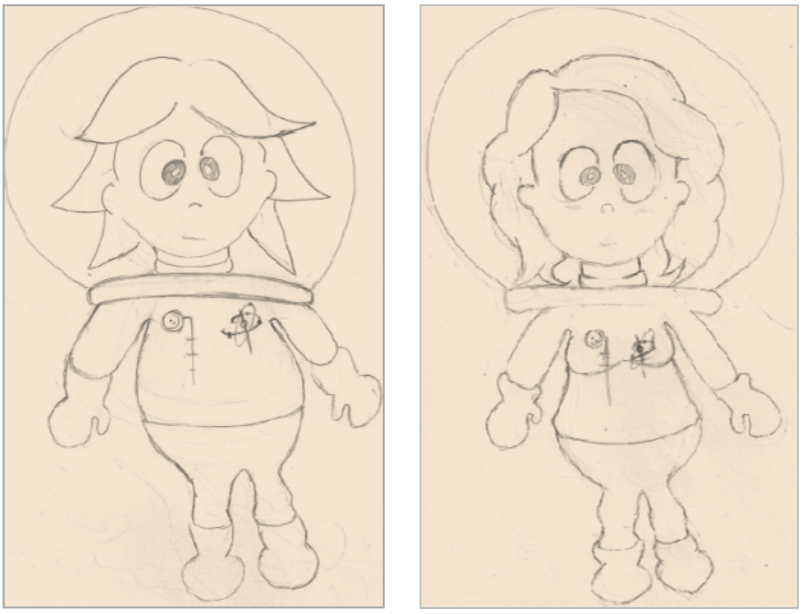


Figure C.1 Pencil drafts of the PlayPhysics player characters

Figure C.2 Player characters texturised in 3D studio Max & animated in Poser

Figure C.3 shows the Python script employed to create the bones and links of the kinematics chain of PlayPhysics player characters in Poser 9. The numbers indicate the level in the chain, followed by the part of the body and configuration of the axes.

```
1 hip yzx :Runtime:Geometries:Spaceman:OBJs:hip.obj
    2 waist yzx :Runtime:Geometries:Spaceman:OBJs:waist.obj
        3 abdomen yzx :Runtime:Geometries:Spaceman:OBJs:abdomen.obj
            4 chest yzx :Runtime:Geometries:Spaceman:OBJs:chest.obj
                5 helmet-bkpack yzx :Runtime:Geometries:Spaceman:OBJs:helmet-bkpack.obj
                5 neck yzx :Runtime:Geometries:Spaceman:OBJs:neck.obj
                    6 head yzx :Runtime:Geometries:Spaceman:OBJs:head.obj
                5 rCollar xyz :Runtime:Geometries:Spaceman:OBJs:rCollar.obj
                    6 rShldr xyz :Runtime:Geometries:Spaceman:OBJs:rShldr.obj
                      7 rForeArm xyz :Runtime:Geometries:Spaceman:OBJs:rForeArm.obj
                        8 rHand xyz :Runtime:Geometries:Spaceman:OBJs:rHand.obj
                         9 rThumb1 xyz :Runtime:Geometries:Spaceman:OBJs:rThumb1.obj
                          10 rThumb2 xyz :Runtime:Geometries:Spaceman:OBJs:rThumb2.obj
                           11 rThumb3 xyz :Runtime:Geometries:Spaceman:OBJs:rThumb3.obj
                         9 rIndex1 xyz :Runtime:Geometries:Spaceman:OBJs:rIndex1.obj
                          10 rIndex2 xyz :Runtime:Geometries:Spaceman:OBJs:rIndex2.obj
                           11 rIndex3 xyz :Runtime:Geometries:Spaceman:OBJs:rIndex3.obj
                5 lCollar xyz :Runtime:Geometries:Spaceman:OBJs:lCollar.obj
                    6 lShldr xyz :Runtime:Geometries:Spaceman:OBJs:lShldr.obj
                      7 lForeArm xyz :Runtime:Geometries:Spaceman:OBJs:lForeArm.obj
                        8 lHand xyz :Runtime:Geometries:Spaceman:OBJs:lHand.obj
                         9 lThumb1 xyz :Runtime:Geometries:Spaceman:OBJs:lThumb1.obj
                          10 lThumb2 xyz :Runtime:Geometries:Spaceman:OBJs:lThumb2.obj
                           11 lThumb3 xyz :Runtime:Geometries:Spaceman:OBJs:lThumb3.obj
                         9 lIndex1 xyz :Runtime:Geometries:Spaceman:OBJs:lIndex1.obj
                          10 lIndex2 xyz :Runtime:Geometries:Spaceman:OBJs:lIndex2.obj
                           11 lIndex3 xyz :Runtime:Geometries:Spaceman:OBJs:lIndex3.obj
    2 rThigh yzx :Runtime:Geometries:Spaceman:OBJs:rThigh.obj
        3 rShin yzx :Runtime:Geometries:Spaceman:OBJs:rShin.obj
            4 rFoot yzx :Runtime:Geometries:Spaceman:OBJs:rFoot.obj
                5 rToe yzx :Runtime:Geometries:Spaceman:OBJs:rToe.obj
    2 lThigh yzx :Runtime:Geometries:Spaceman:OBJs:lThigh.obj
        3 lShin yzx :Runtime:Geometries:Spaceman:OBJs:lShin.obj
            4 lFoot yzx :Runtime:Geometries:Spaceman:OBJs:lFoot.obj
                5 lToe yzx :Runtime:Geometries:Spaceman:OBJs:lToe.obj
ikChain LeftHand lShldr lForeArm lHand
ikChain RightHand rShldr rForeArm rHand
ikChain RightLeg rThigh rShin rFoot
ikChain LeftLeg lThigh lShin lFoot
```

Figure C.3 Specifying the inverse kinematics chain of PlayPhysics player characters

Appendix D. Derivation of PlayPhysics evaluation equations

This section explains the derivation of PlayPhysics equations for evaluating the quality of the response provided by the students to PlayPhysics first game challenge. Also at the end of this section are shown some examples of the non trivial prospective solutions to PlayPhysics first challenge.

Equation D.1 is a powerful expression to help determine the velocity of an object, e.g. Alpha Centauri spaceship, at any time t_s, if we know the object's initial velocity and its constant acceleration (Serway and Jewett 2004). In the first challenge a deceleration must be applied to Alpha Centauri. This is represented by the negative sign before the product 'at' in Equation D.1.

$$v_f = v_i - at_s \qquad \text{Eq. D.1}$$

Also, the final velocity (v_f) should be zero to make Alpha Centauri stop at Athena, after making this substitution and finding t as function of the initial velocity (v_i) and the deceleration (a), the Equation D.2 is obtained.

$$t_s = \frac{v_i}{a} \qquad \text{Eq. D.2}$$

The expression employed for calculating the displacement d_s at constant deceleration is shown in Equation D.3. If t_s is substituted by the derived expression in Equation D.2 to obtain an expression for Equation D.3 as a function of v_i and a is obtained the expression show in Equation D.4.

$$d_s = v_i t - \frac{1}{2} a t_s^{\,2} \qquad \text{Eq. D.3}$$

$$d_s = \frac{v_i^{\,2}}{2a} \qquad \text{Eq. D.4}$$

D.1 Examples of PlayPhysics game challenge prospective solutions

Considering that the maximum magnitude of acceleration is $40\ m/s^2$ (before a person blacks out) and the range of values that can be assigned to the initial velocity, i.e. $[1000, 2000]\ m/s$, here are presented some examples of the values that can be assigned to v_i and a. The cases indicated in blue represent successful cases, i.e. they meet all the game conditions. As can be observed the solution to the scenario is relatively simple when D and T are large, since the combinations of v_i and a that fulfil the requirements are greater. On the contrary, if T and/or D are too small, the valid combinations become scarce.

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
80	70	28	1980	2	68.6	71.4	5	66.5	73.5	70.7	OK
80	70	25	1871	2	68.6	71.4	5	66.5	73.5	74.8	OK
80	70	24	1833	2	68.6	71.4	5	66.5	73.5	76.4	OK
80	70	23	1794	2	68.6	71.4	5	66.5	73.5	78.0	OK
80	70	22	1755	2	68.6	71.4	5	66.5	73.5	79.8	OK
80	70	20	1673	2	68.6	71.4	5	66.5	73.5	83.7	$t_s > T$
80	70	15	1440	2	68.6	71.4	5	66.5	73.5	96.6	$t_s > T$
80	70	10	1183	2	68.6	71.4	5	66.5	73.5	118.3	$t_s > T$

Table D.1 Analysis corresponding to the constraint variables T = 80 s and D = 70 km

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
90	70	25	1871	2	68.6	71.4	5	66.5	73.5	74.8	OK
90	70	20	1673	2	68.6	71.4	5	66.5	73.5	83.7	OK
90	70	15	1449	2	68.6	71.4	5	66.5	73.5	96.6	$t_s > T$
90	70	10	1183	2	68.6	71.4	5	66.5	73.5	118.3	$t_s > T$

Table D.2 Analysis corresponding to the constraint variables T = 90 s and D = 70 km

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
100	70	25	1871	2	68.6	71.4	5	66.5	73.5	74.8	OK
100	70	20	1673	2	68.6	71.4	5	66.5	73.5	83.7	OK
100	70	15	1449	2	68.6	71.4	5	66.5	73.5	96.6	OK
100	70	10	1183	2	68.6	71.4	5	66.5	73.5	118.3	$t_s > T$

Table D.3 Analysis corresponding to the constraint variables T = 100 s and D = 70 km

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
110	70	25	1871	2	68.6	71.4	5	66.5	73.5	74.8	OK
110	70	20	1673	2	68.6	71.4	5	66.5	73.5	83.7	OK
110	70	15	1449	2	68.6	71.4	5	66.5	73.5	96.6	OK
110	70	10	1183	2	68.6	71.4	5	66.5	73.5	118.3	$t_s > T$

Table D.4 Analysis corresponding to the constraint variables T = 110 s and D = 70 km

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
120	70	25	1871	2	68.6	71.4	5	66.5	73.5	74.8	OK
120	70	20	1673	2	68.6	71.4	5	66.5	73.5	83.7	OK
120	70	15	1449	2	68.6	71.4	5	66.5	73.5	96.6	OK
120	70	10	1183	2	68.6	71.4	5	66.5	73.5	118.3	OK

Table D.5 Analysis corresponding to the constraint variables T = 110 s and D = 70 km

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
80	60	30	1897	2	58.8	61.2	5	57	63	63.2	OK
80	60	25	1732	2	58.8	61.2	5	57	63	69.3	OK
80	60	20	1549	2	58.8	61.2	5	57	63	77.5	OK
80	60	15	1342	2	58.8	61.2	5	57	63	89.4	$t_s > T$
80	60	10	1095	2	58.8	61.2	5	57	63	109.75	$t_s > T$

Table D.6 Analysis corresponding to the constraint variables T = 80 s and D = 60 km

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
80	50	40	2000	2	49	51	5	47.5	52.5	50	OK
80	50	35	1871	2	49	51	5	47.5	52.5	53.5	OK
80	50	30	1732	2	49	51	5	47.5	52.5	57.7	OK
80	50	25	1581	2	49	51	5	47.5	52.5	63.2	OK
80	50	20	1414	2	49	51	5	47.5	52.5	70.7	OK
80	50	15	1225	2	49	51	5	47.5	52.5	81.6	$t_s > T$
80	50	10	1000	2	49	51	5	47.5	52.5	100	$t_s > T$

Table D.7 Analysis corresponding to the constraint variables T = 80 s and D = 50 km

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
80	40	50	2000	2	39.2	40.8	5	38	42	40	$a = 40\,m/s^2$
80	40	40	1789	2	39.2	40.8	5	38	42	44.7	OK
80	40	35	1673	2	39.2	40.8	5	38	42	47.8	OK
80	40	30	1549	2	39.2	40.8	5	38	42	51.6	OK
80	40	25	1414	2	39.2	40.8	5	38	42	56.6	OK
80	40	20	1265	2	39.2	40.8	5	38	42	63.2	OK
80	40	15	1095	2	39.2	40.8	5	38	42	73	OK

Table D.8 Analysis corresponding to the constraint variables T = 80 s and D = 40 km

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
80	30	50	2000	2	29.4	30.6	5	28.5	31.5	34.6	$a = 40\,m/s^2$
80	30	40	1789	2	29.4	30.6	5	28.5	31.5	38.7	OK
80	30	35	1673	2	29.4	30.6	5	28.5	31.5	41.4	OK
80	30	30	1549	2	29.4	30.6	5	28.5	31.5	44.7	OK
80	30	25	1414	2	29.4	30.6	5	28.5	31.5	49	OK
80	30	20	1265	2	29.4	30.6	5	28.5	31.5	54.8	OK

Table D.9 Analysis corresponding to the constraint variables T = 80 s and D = 30 km

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
80	20	50	1414	2	19.6	20.4	5	19	21	28.3	$a = 40\,m/s^2$
80	20	40	1265	2	19.6	20.4	5	19	21	31.6	OK
80	20	35	1183	2	19.6	20.4	5	19	21	33.8	OK
80	20	30	1095	2	19.6	20.4	5	19	21	36.5	OK
80	20	25	1000	2	19.6	20.4	5	19	21	40	OK

Table D.10 Analysis corresponding to the constraint variables T = 80 s and D = 20 km

$T[s]$	$D[Km]$	$a[m/s^2]$	$v_i[m/s]$	$e_{d_1}[\%]$	$d_{1_{min}}$	$d_{1_{max}}$	$e_{d_2}[\%]$	$d_{2_{min}}$	$d_{2_{max}}$	$t_s[s]$	Result
80	15	50	1225	2	14.7	15.3	5	14.25	15.75	28.3	$a = 40\,m/s^2$
80	15	40	1095	2	19.6	20.4	5	19	21	31.6	OK
80	15	35	1025	2	19.6	20.4	5	19	21	33.8	OK

Table D.11 Analysis corresponding to the constraint variables T = 80 s and D = 15 km

Appendix E. PlayPhysics software implementation

This section presents complementary code sections related to PlayPhysics functioning.

E.1 Code fragments related to PlayPhysics communication

```
function Awake()
{
        stdobj = student.GetComponent(StudentData);
        if(stdobj)
        {
                stdobj.Disarm();
                Application.ExternalCall("submitDataUnity", " ");
        }
        else
        {
                Debug.Log("Add the student object");
        }
}
```

Figure E.1 Awake' method as implemented in StartMenuGUI.js

```
function submitDataUnity()
                {
                        var user = document.form1.usernick.value;
                        var last = document.form1.lastnick.value;
                        var sex = document.form1.sexnick.value;
                        var gprogress = document.form1.gameprogress.value;
                        var url = document.form1.servurl.value;

                        var unityParams =user + "," + last  + "," + sex  + "," + gprogress ","
                        +url;
                        GetUnity().SendMessage("Main Camera", "SetStudentVars",
                        unityParams);
                }
```

Figure E.2 Code fragment corresponding to the 'submitDataUnity' function

```
function loadXMLFile()
{
        if(stdobj)
        {
        //Load the XML file
        var URL =
            stdobj.GetStudentServerURL()+"/PlayPhysics/jsp/student/unitycontent/dialogueFirstChallenge.xml";
        var xmlFile = new WWW(URL);
        yield xmlFile;
        if (xmlFile.error == null)//If there is no error
```

```
                {

                        tXMLParser = new XMLParser();
                        tDialogueData= tXMLParser.ParseString(xmlFile.text);
                        AnswersToQuestions = new Array();
                }
        else
                {
                        Debug.Log("It was an error loading the xml file:" + xmlFile.error.ToString());
                }
        }
  }
```

Figure E.3 Code fragment employed for loading an XML file in PlayPhysics

```
function saveSession()
{
        if(AnswersToQuestions!=null && stdobj!=null)
        {
                var form = new WWWForm();
...
                form.AddField("username", stdobj.GetStudentID());
                //Attitude Beliefs towards physics
                form.AddField("attitudetophysics",  attitudetophysics);
                //Confidence: Attitude towards success
                form.AddField("confidence",  confidence);
                //Source of motivation
                form.AddField("sourcemotivation",  sourcemotivation);
                //Perceived level of difficulty
                form.AddField("perceiveddifficulty",  perceiveddifficulty);
                //Attitude towards Effort
                form.AddField("attitudetoeffort",  attitudetoeffort);
                form.AddField("emotion",  emotion);
                // Go to the tutorial
                form.AddField("gameProgress", "2");
                var URL =
                stdobj.GetStudentServerURL()+"/PlayPhysics/student/RegisterFirstDialogue";
                var w  = new WWW(URL, form);
                debugStr ="Connecting with the server...";
                errorFlag="true";
                yield WaitForRequest(w);

                if(tResponseData!=null)
                {       if(tResponseData[0][1]!=null)
                        {
                                if(tResponseData[0][1]["name"] =="result")
                                saveResult =tResponseData[0][1]["value"];
                        }
                }
```

```
        }
}
```

Figure E.4 Fragment of ‘SaveSession’ function using WWW to send data

E.2 Code fragment related to PlayPhysics simulation and world models

```
function InitAll()
{
...
        PhaseIGUIObj = GameObject.Find("AlphaCentauriShip");
        if(PhaseIGUIObj)
        {       //Challenge variables
                SetValInitialDistance(UnitConversion.ToMeters(initialDistance), true);
                SetValTimeFinishCombustible(timeLimit);

                //Variables of interaction for the student
                PhaseIGUIObj.GetComponent("FirstLevelGUI_PhaseI").valInitialVelocity = ""+ (UnitConver-
                sion.ToMeters(MAX_INITIAL_VELOCITY));
                PhaseIGUIObj.GetComponent("FirstLevelGUI_PhaseI").valAcceleration = "" + (UnitConver-
                sion.ToMeters(MIN_FWD_BWD_ACCELERATION));
                PhaseIGUIObj.GetComponent("FirstLevelGUI_PhaseI").proportionDistanceAthena = (initialDis-
                tance/initialDistanceToBoundary);
                if (reStartFlag)
                {
                        PhaseIGUIObj.GetComponent("FirstLevelGUI_PhaseI").introFlag = false;
                }

        }
        else
                Debug.Log("ChallengeObject_PhaseI: FirstLevelGUI_PhaseI does not exist");

        if(stdobj && PhaseIGUIObj)//The display of PlayPhysics game challenge GUI is handled
        {
                //stdobj.
                //stdobj.sendDuringInteractionData();//Validate that the insert was ok
        }
}
```

Figure E.5 Code fragment from ChallengeController.js

Appendix F. Alpha Centauri spaceship's functionality design

This section presents the initial pencil drafts of the Alpha Centauri spaceship. As can observed from Figures F.1 and F.2, the spaceship was designed to perform displacement backwards and forwards in three axes: ö, Ê and û, and perform three rotations: roll, pitch, yaw.

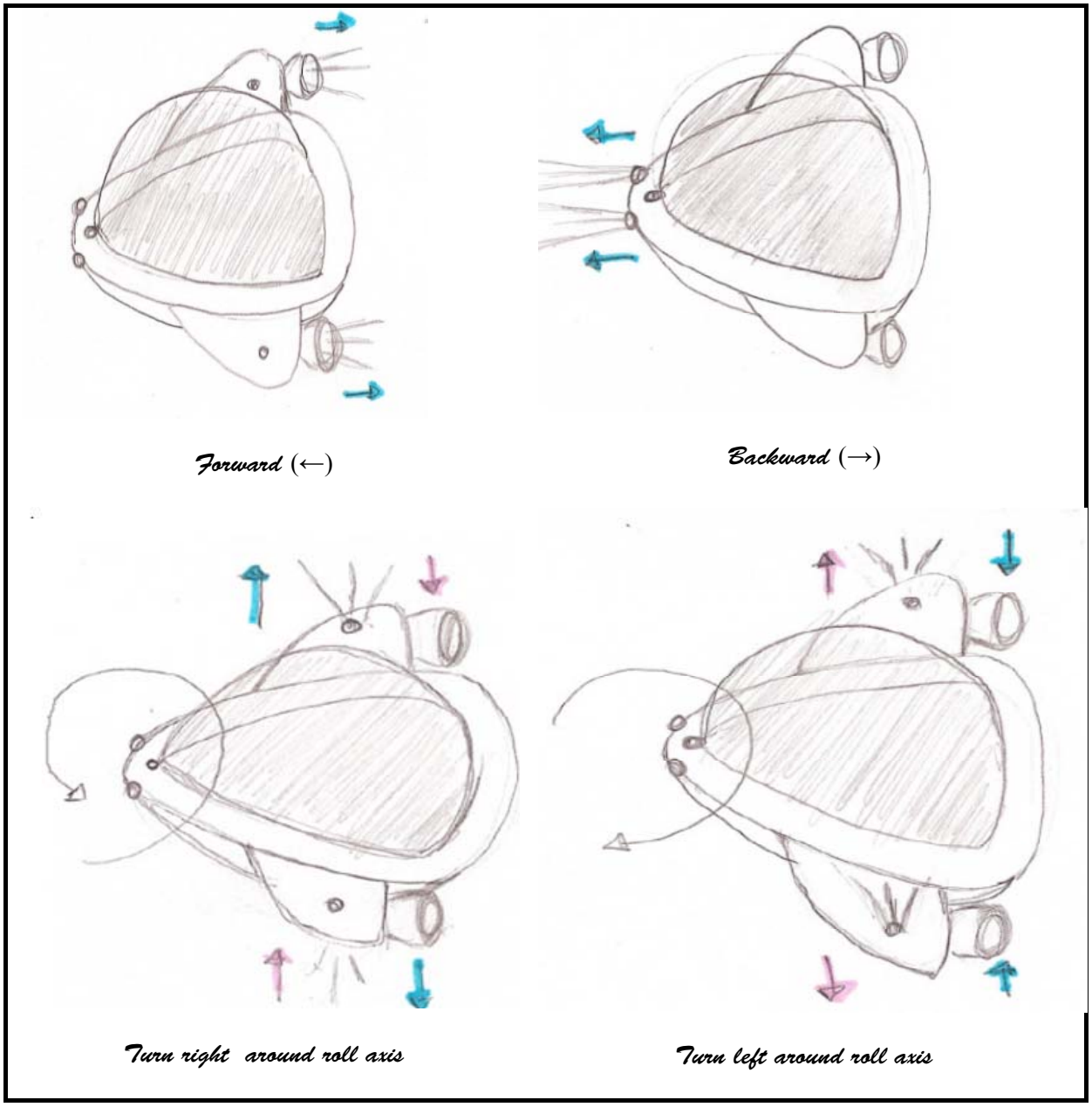


Figure F.1 Alpha Centauri's displacements and its rotations around the roll axis

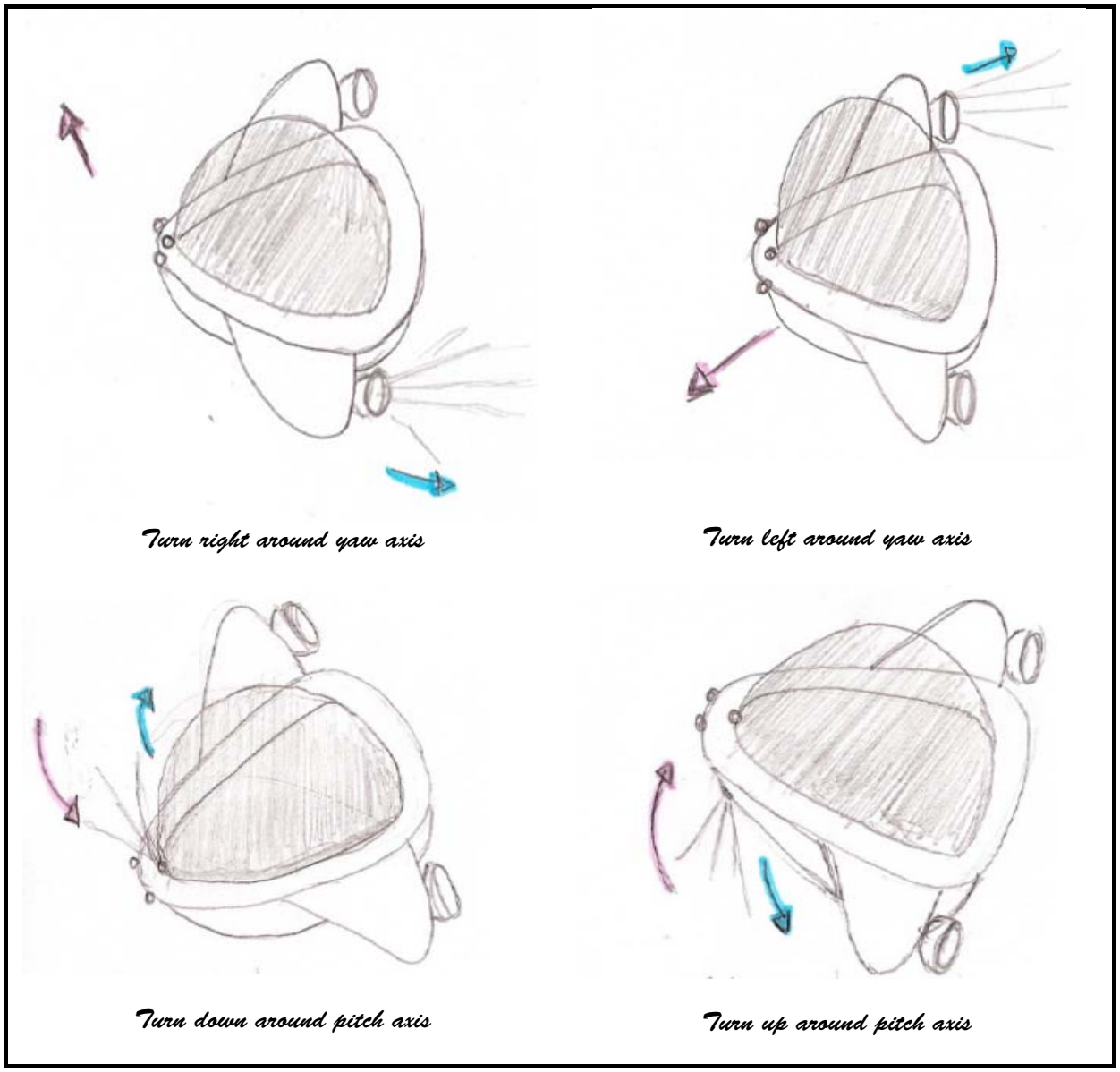


Figure F.2 Alpha Centauri rotations around yaw and pitch axes

Appendix G. PlayPhysics tutorials

Two types of materials were created for students to support their interaction with PlayPhysics: (1) an installation manual and (2) an interaction tutorial. The former serves the purpose of aiding students to install all the software required to play PlayPhysics: a) Mozilla Firefox, version 6.0.1 or over and b) the Unity3D web player. The latter introduces to the student the functionality of the game GUI.

The Mozilla Firefox browser is required for the appropriate operation of PlayPhysics, since the JavaScript code designed for it. At the beginning, the installation tutorial was created as a word file with screenshots describing step by step how the installation has to be performed. However, it was observed that students were having several problems following the document. As a result, lecturers suggested making this tutorial as a video using Camtasia (TechSmith 2011) and uploading it to YouTube. It was included as a link in PlayPhysics' main web page, see Figure G.1. Clicking the link will open the video in the YouTube page (Figure G.2). After deploying the video, no further problems related to following the steps of installation were reported.

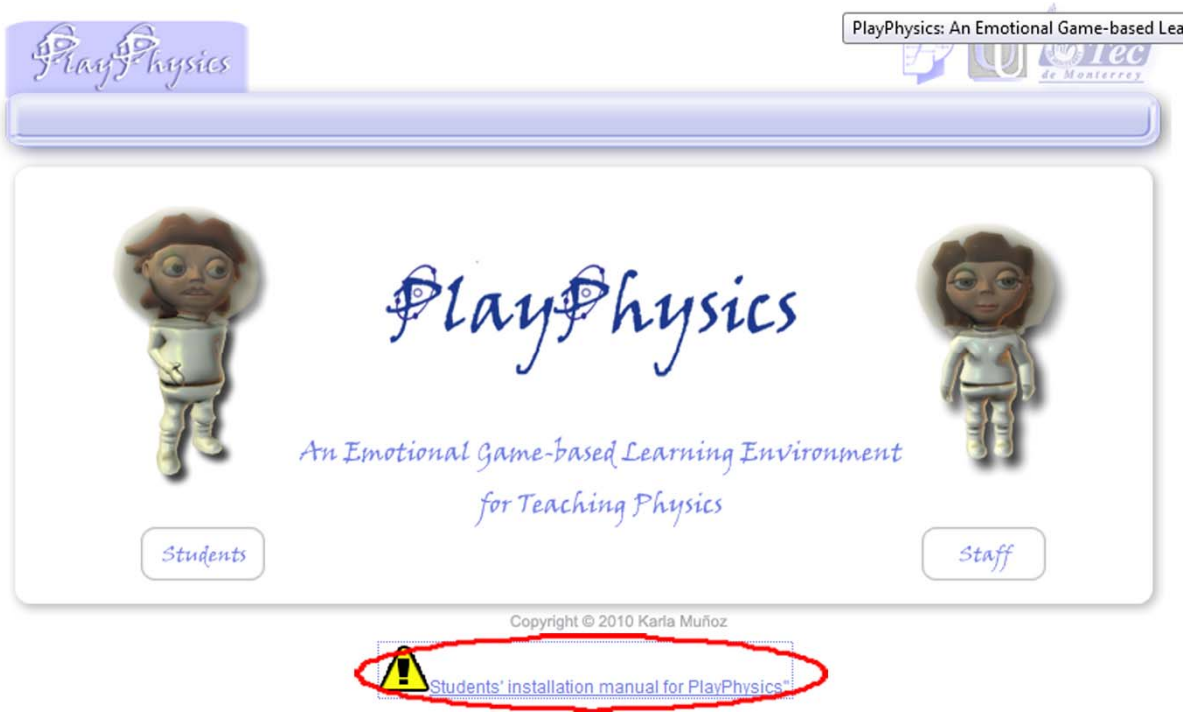


Figure G.1 Link to PlayPhysics installation manual in PlayPhysics main page

The interaction tutorial was incorporated as part of PlayPhysics educational game flow. It is displayed just before the student proceeds to interact with PlayPhysics game challenges and the student can go through it at his/her own pace. In this tutorial M8 introduces himself, the goals of the game and the functionality of its interface, see Figures G.3 to G.6.

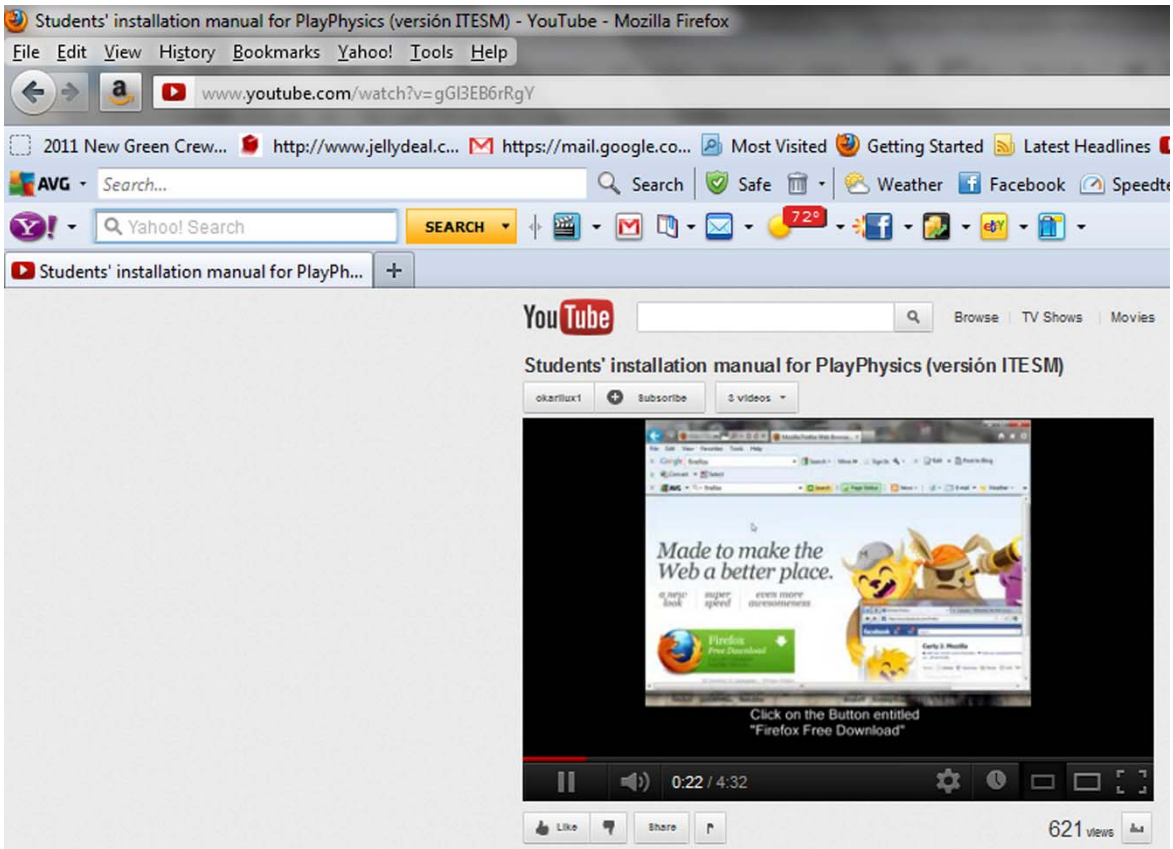


Figure G.2 PlayPhysics installation manual opened in YouTube

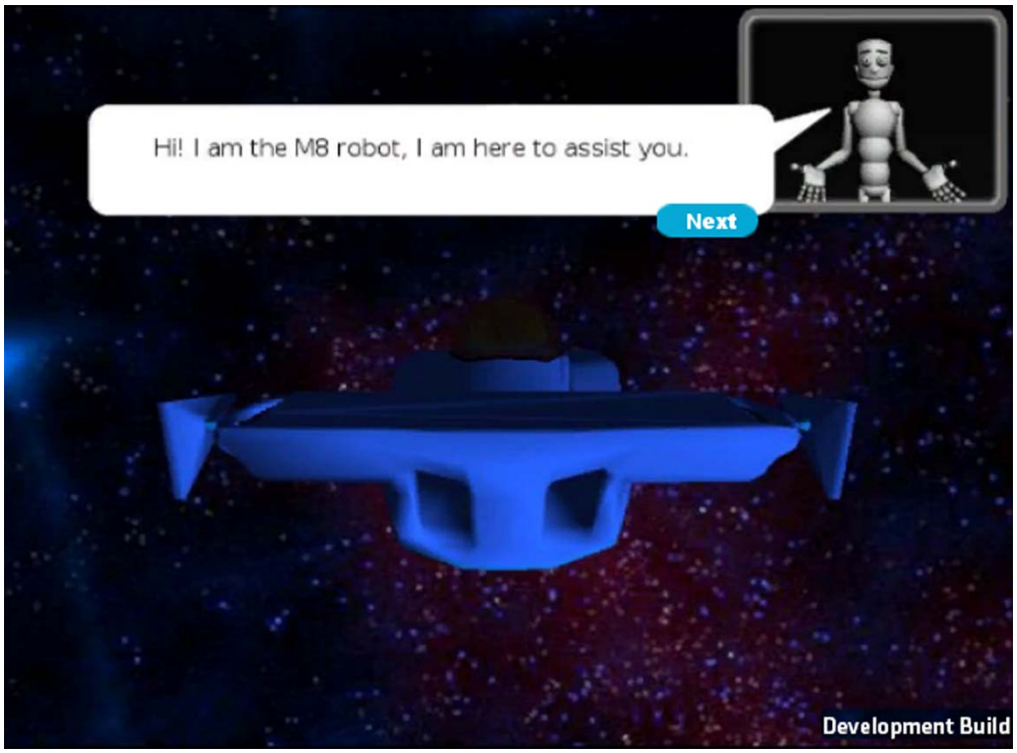


Figure G.3 Screenshot illustrating how M8 introduces itself

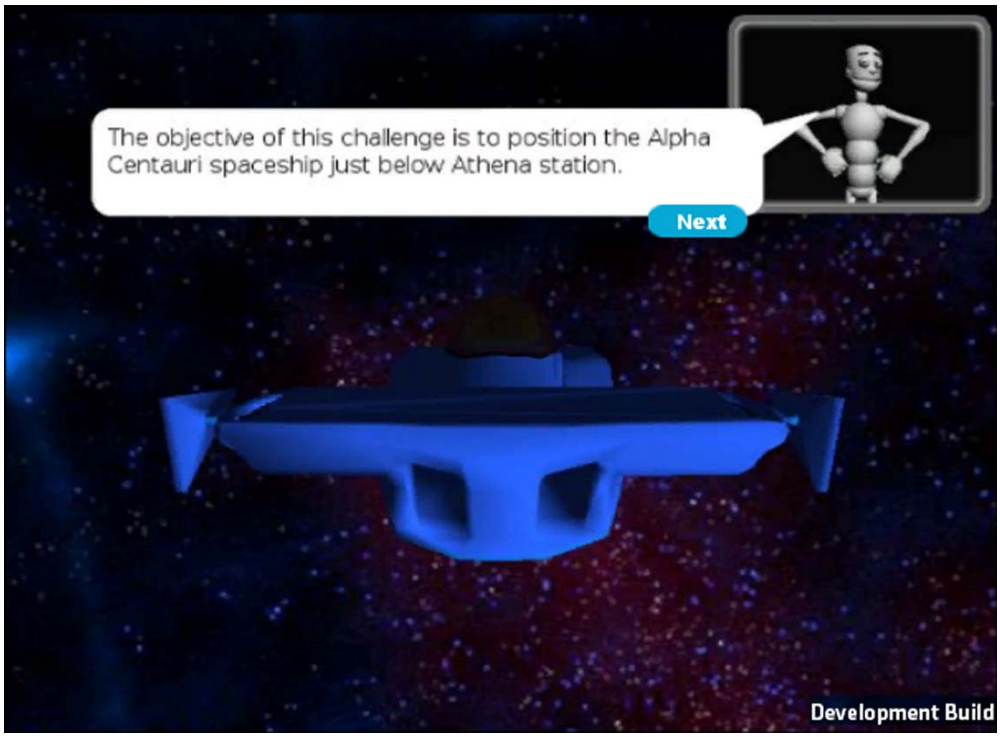


Figure G.4 Screen shot of M8 introducing the game challenge goals

The explanation about the game functionality starts by explaining how to change between the available views (Figure G.5) and how to set the values of the exploration variables: initial velocity and acceleration (Figure G.6).

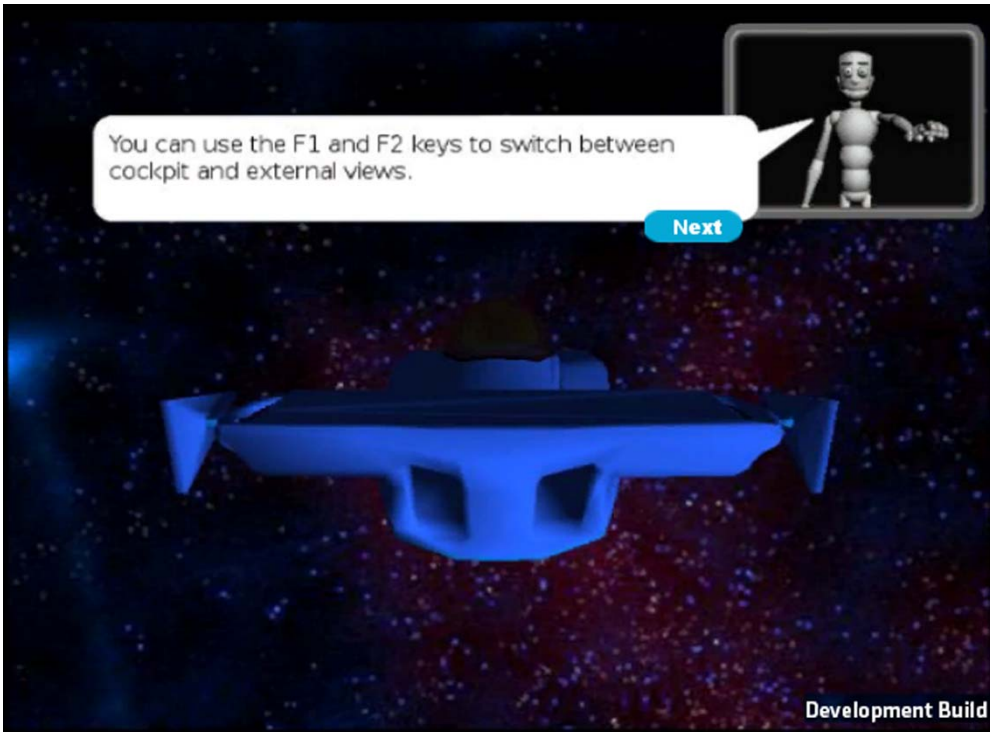


Figure G.5 Screen shot of M8 introducing PlayPhysics game challenge functionality

M8 then teaches the student how to report her emotion and lets her know that it will occasionally ask the student to report his/her emotion, see Figure G.7. In addition, M8 explains to the student how to ask for a hint in case that she thinks that it is necessary.

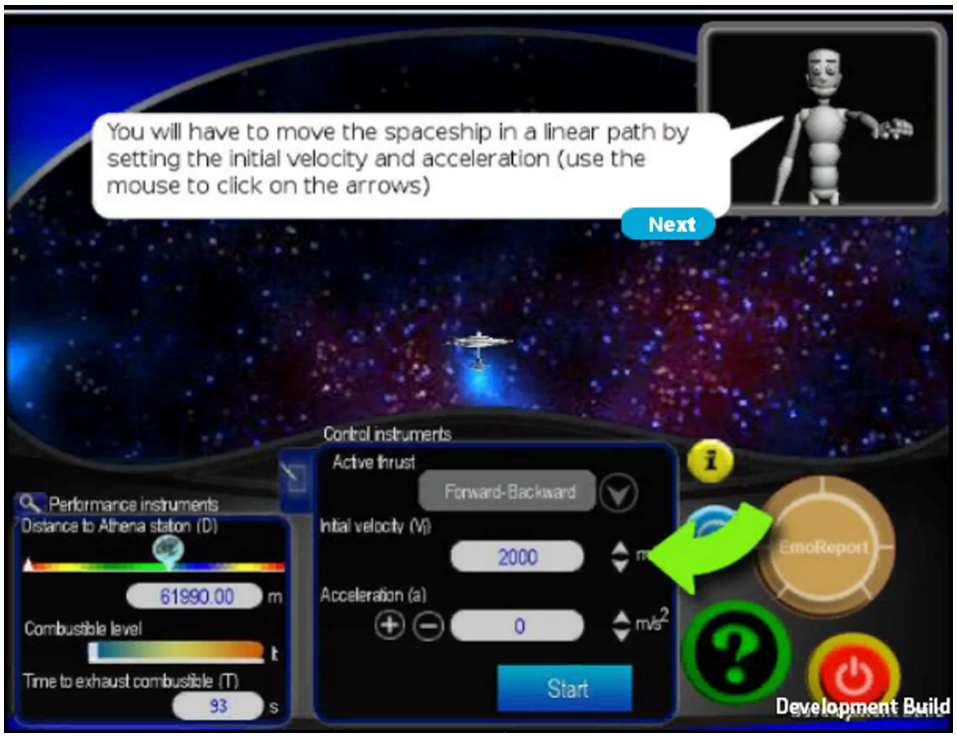


Figure G.6 Screen shot of M8 showing the student how to achieve the game goals

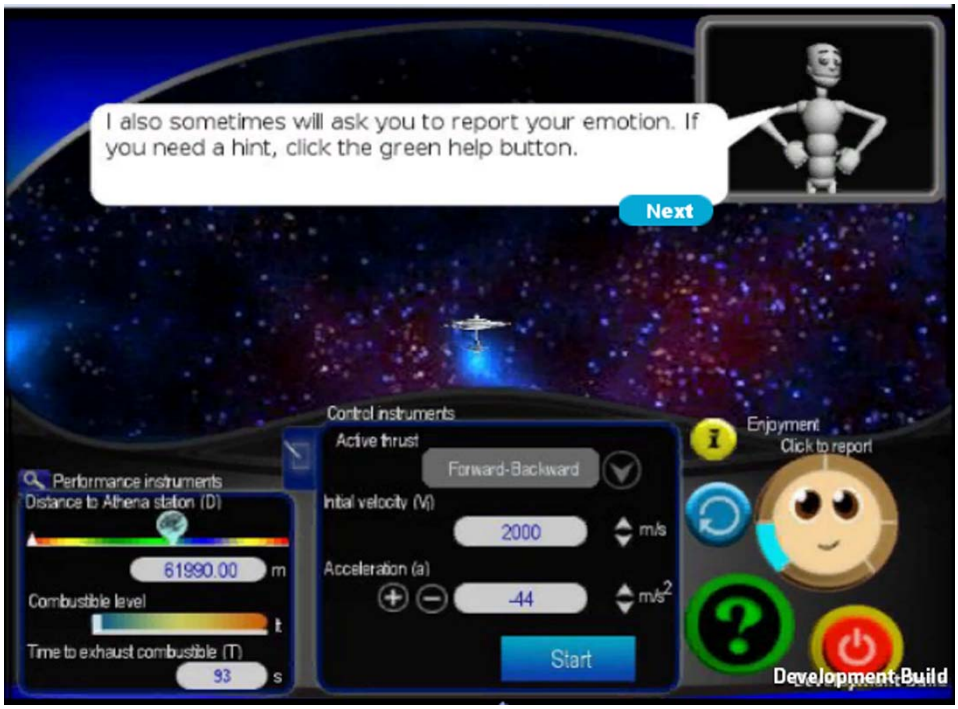


Figure G.7 Screenshot where M8 explains the student about self-reporting emotions

Appendix H. PlayPhysics pre-test and post-test

This section shows the pre-test and post-test used to assess the student knowledge and understanding about the physics concepts taught by the PlayPhysics educational game and the concepts presentation. Both tests were reviewed by an expert lecturer and Astrophysics at ITESM-CCM before been implemented in PlayPhysics. Students in the focus group will also have to answer a qualitative questionnaire, which was created to understand how the students perceive PlayPhysics' educational game. This qualitative questionnaire is also shown at the end of this section.

Pre-test

Please select the correct answer to each question:

1) Which of the following statements is true in a one-dimensional motion when acceleration is constant?

- a. The average acceleration over any time interval is different to the instantaneous acceleration at any instant within that interval
- b. The velocity changes are not-uniform overtime
- c. The acceleration-time graph is a straight line with a slope equal to zero
- d. The one-dimensional motion is more complex and difficult to analyse than a motion where acceleration is not constant

2) Considering a positive axis to the right, if a car slows down as it moves to the right. It is because ...

- a. the car has a negative acceleration
- b. its displacement increases with time
- c. the magnitude of the velocity vector increases with time
- d. the velocity and acceleration vectors have the same direction

Car Breaking

Car A is waiting at a crossing for the green light, while car B is approaching car A from behind. When car B is 45 m from car A, what should its initial velocity and constant acceleration be to

stop within 3 seconds just before hitting car A?

3) The car's initial velocity is:

- a. 90 m/s
- b. 40 km/h
- c. 30 m/s
- d. 60 km/h

4) the car's acceleration should be:

- a. 8 m/s^2
- b. -21 km/h^2
- c. -10 m/s^2
- d. 3 m/s^2

Return to the main menu

Post-test

Please select the correct answer to each question:

Jet Landing

If a jet advances much farther than 63 m while landing on an aircraft carrier, it might fall into the ocean. Calculate its initial velocity and acceleration if its landing lasts 2 seconds?

1) The jet's initial velocity must be:

- a. 140 m/s
- b. 100 km/h
- c. 80 km/h
- d. 63 m/s

2) The jet's acceleration must be:

- a. 28 m/s^2
- b. -31 m/s^2
- c. 11 m/s^2
- d. -18m/s^2

3) Which of the following statements is true in a one-dimensional motion when acceleration is constant?

- a. The average acceleration over any time interval is equal to the instantaneous acceleration at any instant within that interval
- b. The slope of a displacement-time graph of an object in movement is a straight line and describes its acceleration
- c. The velocity-time graph is a straight-line with a slope of zero
- d. The rate of change of the velocity varies randomly throughout the motion

4) If a car moves to the right increasing its displacement between adjacent positions over time, it is because...

- a. The car has a negative acceleration to the left
- b. Its acceleration increases with time
- c. The velocity vector decreases with time
- d. The velocity and acceleration vectors are in the same direction

Return to the main menu

Qualitative Evaluation of PlayPhysics

Thank you for helping us to test PlayPhysics. This questionnaire was created to know about your perception of PlayPhysics. Please assist us answering it

Fantasy and Curiosity

1) The story of the game is appealing

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

2) PlayPhsyics encourage you to replay and increase your score

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

3) Interacting with PlayPhysics encourages you to continue interacting with it and to solve other challenges

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

4) The game challenge about arriving to the space station encourage you to think on other situations where you have to apply the same physics principles to face a similar challenge

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

Challenge

5) The goal of the game is explicit

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

6) The feedback (sounds, graphics and final score) provided is appropriate and lets you know how close you are to achieving the challenge

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

7) The knowledge or feedback provided by the M8 robot is appropriate and helps you understand the physics principles behind the game challenge

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

8) The challenge is achievable

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

9) PlayPhysics assists you in enhancing your understanding of the related physics principles and concepts

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

Ease of interaction

10) The Graphical User Interface (GUI) is intuitive

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

M8 Behaviour

11) The emotional behaviour of the robot M8 is appropriate

☐ Highly disagree ☐ Disagree ☐ Neither Agree nor Disagree ☐ Agree ☐ Highly Agree

Suggestions for improvements

(Maximum characters: 240)

You have 240 characters left.

Additional Comments

(Maximum characters: 240)

You have 240 characters left.

Return to the main menu

Appendix I. Subject consent form

The ethical approval at ITESM-CCM did not require a formal application or consent form for the interaction of students with PlayPhysics via the web, but it required lecturer pre-testing of PlayPhysics and authorisation for proceeding to test with students. However, for the interaction of PlayPhysics on-site to experimenting with the GSR Bluetooth sensor, a student consent form was required, which was required by ITESM-CCM. This form is following presented:

Hoja Informativa del Participante

University of Ulster, Magee campus,
Room MS125
Intelligent Systems Research Centre,
BT48 7JL, Co. Derry/Londonderry,
Northern Ireland, UK

Tecnológico de Monterrey (ITESM),
Mexico City campus,
Escuela de Ingeniería y Arquitectura
Calle del Puente 222
Col. Ejidos de Huipulco Tlalpan,
C.P. 14380, México, D.F.

FORMA DE CONSENTIMIENTO INFORMADO

Nombre del proyecto: PlayPhysics: An Emotional Student Model for Intelligent Gaming
Versión 1, 15th September 2011
Investigadores: Karla Muñoz**,** Prof. Paul Mc Kevitt, Dr. Tom Lunney, Dr. Julieta Noguez, Dr. Luis Neri
Investigador de contacto en Ulster: Karla Muñoz, munoz_esquivel-k@email.ulster.ac.uk
Investigador de contacto en ITESM: Julieta Noguez, jnoguez@itesm.mx

Esta investigación busca hacer a las computadoras capaces de reconocer y entender las emociones de estudiantes que se encuentran persiguiendo un título a nivel universitario en Ciencias o Ingeniería y en un rango de edad de 18 a 23 años. El objetivo final es crear ambientes de aprendizaje basado en juegos que puedan adaptarse al nivel de entendimiento y motivación de cada estudiante. Si decides participar en nuestra investigación, tendrás que usar un sensor usado para medir la respuesta Galvánica de la piel en dos dedos correspondientes a tu mano izquierda (diestro/a) o derecha (zurdo/a) mientras interactúas con PlayPhysics y reportar tu emoción usando los botones diseñados y nombrados "*EmoReport"*. Este sensor mide la resistencia de tu piel, ya que previa investigación ha demostrado que la variación de la resistencia está relacionada con los cambios emocionales.

Para medirla, el sensor tiene que aplicar una pequeña corriente a tu piel. En algunos casos, ésta pequeña corriente puede causar irritación en tu piel. Por lo que para evitar este problema estamos usando un gel especial que previene este efecto, sin embargo algunos tipos de piel son más delicados que otros, por lo que si experimentas comezón entre tus dedos mientras te encuentras usando el sensor, por favor notifica al investigador a cargo del experimento para que pare el experimento.

1. Confirmo que he leído y entendido la información de la "Hoja Informativa del Participante" con fecha del 15 de Septiembre del 2011, versión 1, para la investigación previamente descrita y tuve oportunidad de hacer todas las preguntas necesarias y las preguntas acerca de los posibles riesgos y beneficios han sido respondidas completamente..... ☐
2. Entiendo que mi participación es voluntaria y soy libre de abandonar esta investigación a cualquier momento sin dar razón alguna y sin que mis derechos legales sean afectados.... ☐
3. Entiendo que la información recolectada como parte de esta investigación, tal vez será revisada por personal responsable de la University of Ulster e ITESM o por autoridades regulatorias para las cuales es relevante mi participación en esta investigación. Otorgo permiso a estas personas para acceder a mi información, la cual es relevante para esta investigación..... ☐

Nombre: _____ Matrícula: _____

Signature: _____

Subject information sheet

University of Ulster, Magee campus,
Room MS125
Intelligent Systems Research Centre,
BT48 7JL, Co. Derry/Londonderry,
Northern Ireland, UK

Tecnológico de Monterrey (ITESM),
Mexico City campus,
Escuela de Ingeniería y Arquitectura
Calle del Puente 222
Col. Ejidos de Huipulco Tlalpan,
C.P. 14380, México, D.F.

INFORMED CONSENT FORM

Name of the project: PlayPhysics: An Emotional Student Model for Intelligent Gaming
Version 1, 15th September 2011
Researchers: Karla Muñoz**,** Prof. Paul Mc Kevitt, Dr. Tom Lunney, Dr. Julieta Noguez, Dr. Luis Neri
Contact investigator at University of Ulster: Karla Muñoz, munoz_esquivel-k@email.ulster.ac.uk
Contact investigator at ITESM: Julieta Noguez, jnoguez@itesm.mx

This research aims to enable computers to recognise students' emotions, aged between 18 and 23 and willing to undertake or undertaking education in Science or Engineering related to the study of physics. The ultimate goal is to create game-based learning environments that adapt to the specific level of understanding and engagement of each student. If you decide to participate in our study, you will have to wear a Galvanic Skin Response (GSR) sensor in two fingers of your left hand while interacting with PlayPhysics and reporting your emotion using the *EmoReport* wheel. The sensor measures the resistance of your skin overtime, which varies depending on your emotional state. To measure it, the sensor has to pass a small current through your skin. In some cases this current can cause a small irritation of the skin. However to avoid this problem, we are employing a special gel, which will protect you of this effect. But if this happens and you start to feel itching, I will ask you to please let the person in charge of the experiment know about this to stop the experiment.

4. I confirm that I have read and understand the subject information sheet dated 15th September 2011, version 1 for the above study and have had the opportunity to ask questions which have been answered fully with respect of possible risk and potential benefits..... ☐
5. I understand that my participation is voluntary and I am free to withdraw at any time without giving any reason, without my legal rights being affected..... ..... ☐
6. I understand that my data collected as part of this research may be looked by responsible individuals from the University of Ulster and ITESM or from regulatory authorities where it is relevant to me taking part in this research. I give permission to these individuals to access this data, which is relevant to this research..... ☐

Name: _____ Registration number: _____

Signature: _____

Appendix J. Modifications of PlayPhysics GBL environment

From the test conducted with PlayPhysics from January to March 2011, four lecturers and eight lecturers gave us suggestions and asked for some changes in order to enhance the interaction with PlayPhysics GBL environment. Students and lecturers suggestions or requirements are summarised in Tables J.1 and J.2. These suggestions were classified into High, Medium, Low and Low/Reject priorities using as criteria the person that asked for the change, whether the suggestion were related to students' performance and the frequency with which the suggestion was reported. If the student and the lecturer asked for the change was considered of higher priority than if only the lecturer asked for the change. Also, in cases where only the students asked for the changes, these were assigned less priority than the previous mentioned cases. Additionally, comments and suggestions directed to enhance the game looks had less priority than issues that impacted performance and functionality.

Suggestion/Requirement	*Reported frequency*	*Suggested/Asked by*	*Priority/Status*
Transparency on game rules: specifically what they have to do and how they are evaluated (marking scheme)	4	Students and lecturers	High
Instead of explaining the astronaut's cause of dead. Make PlayPhysics to give a more implicit explanation, e.g. may be using graphics. (So students that require more help, still have to ask M8 for a hint with a more detailed description of the problem) For example, display the result showing to the student that the player character fainted	1	Lecturers	Medium
Introduce more information about how students can interact with their environment or provide a tutorial: -Show a legend that says use F1 and F2 keys to change between the cockpit and external views -Explain to the student that he/she has to set the values for the initial velocity and acceleration	1	Lecturers	Medium
Display just integer numbers for the distance to Athena space station on the cockpit view	1	Lecturers	Medium
Put the camera in a position where Alpha Centauri spaceship looks smaller than Athena space station	1	Lecturers	Low
Once that the spaceship arrives to Athena space station leave it in that position for a few seconds to make the student understand that he or she achieved the learning goal and then congratulate him/her	2	Lecturers	Medium
Show a negative distance when the spaceship passes the space station	1	Lecturers	Low
Double check that the pre-test and post-tests are functioning	1	Lecturers	Medium
Make the user manual more dynamic and short (probably make a video)	2	Lecturers and students	Medium
Include a way of sending feedback or comments about PlayPhysics interaction by students to the System administrator	1	Lecturers	Low

Table J.1 Summary of changes suggested by students and lecturers part 1

Suggestion/Requirement	*Reported frequency*	*Suggested/Asked by*	*Priority/Status*
About the game challenge music: if possible being able to disable the music or change it	2	Students	Low
Include more missions	1	Students	Low
Incorporate more dynamic ways of controlling (different from keys and mouse)	3	Students	Low
Explain more about the controls on the GUI that are employed to interact, such as Re-start and Quit	3	Lecturers and students	High
Do not put too much text in the game (Introduction and instructions)	1	Students	High
Make the introduction smaller and dynamic	1	Students	High
The game scene should make the student feel that he/she advances when travelling the spaceship. Therefore include asteroids or starts, which are travelling, to create that sensation	2	Students	High
Performance issues: Look how the loading speed can be increased (may be changing the application to other server)	1	Students	High
It would be desired to be able of personalising the player character	1	Students	Low/Rejected
If possible change the player character for one more mature looking	1	Students	Low/Rejected
If feasible change the spaceship for one more realistic looking	1	Students	Low/Rejected
Allow setting values without limits for the velocity and the acceleration	1	Students	Low/Rejected

Table J.2 Summary of changes suggested by students and lecturers part 2

Appendix K. Results

This section presents additional material related to the PlayPhysics' emotional student model or PlayPhysics GBL environment evaluation.

K.1 CPTs of prospective-outcome emotions network

Value	*Pre-test*	
	"(-inf-87.5]"	"(87.5-inf)"
Positive	0.452	0.840
Negative	0.548	0.160

Pre-test	*Control*	
	High	Medium
"(-inf-87.5]"	0.525	0.404
"(87.5-inf)"	0.475	0.596

Table K.1 CPTs corresponding to *value* and *pre-test*

Control	
High	0.435
Medium	0.565

Prospective level of difficulty	
Low/High	0.380
Medium	0.619

Attitude towards effort	
Positive	0.826
None/Negative	0.174

Table K.2 CPTs of *control*, *prospective level of difficulty* and *attitude towards effort*

Source of motivation	*Attitude towards physics*		
	Positive	Negative	None
Inner	0.800	0.250	0.595
External/Both	0.200	0.750	0.405

Attitude towards physics	*Control*	
	High	Medium
Positive	0.500	0.288
Negative	0.250	0.192
None	0.250	0.519

Table K.3 CPTs corresponding to *source of motivation* and *attitude towards physics*

Control		High		Medium	
Value		Positive	Negative	Positive	Negative
Emotion	Anticipatory Joy	1.000	0.000	0.000	0..000
	Hope	0.000	0.000	1.000	0.000
	Anticipatory Relief	0.000	1.000	0.000	0.000
	Anxiety	0.000	0.000	0.000	1.000

Table K.4 CPT corresponding to the *prospective outcome* emotions

Attitude towards physics		Positive				None/High			
Control		High		Medium		High		Medium	
Prospective level of difficulty		Low/High	Medium	Low/High	Medium	Low/High	Medium	Low/High	Medium
Confidence	High	0.667	0.714	0.315	0.481	0.000	0.500	0.000	0.200
	Medium	0.250	0.286	0.625	0.519	1.000	0.250	0.000	0.800
	Low	0.083	0.000	0.063	0.000	0.000	0.250	1.000	0.000

Table K.5 CPT corresponding to *confidence*

K.2 Pearson correlations of prospective outcome emotions variables

The Pearson correlations for the variables involved in the creation of the model corresponding to the prospective outcome emotions are presented for 92 cases.

	Gender	Pre-test	Attitude towards physics	Confidence	Source of motivation	Prospective level of difficulty	Attitude towards effort	Emotion	Value	Control
Gender										
Pearson correlations	1	0.043	0.191	-0.332**	0.172	-0.026	0.255*	-0.195	-0.006	-0.153
Sig. (2-tailed)		0.687	0.068	0.001	0.100	0.807	0.014	0.063	0.957	0.145
Pre-test										
Pearson correlations	0.043	1	0.227*	-0.213*	0.316**	0.136	0.270**	0.114	0.408**	0.121
Sig. (2-tailed)	0.687		0.030	0.042	0.002	0.197	0.009	0.278	0.000	0.252
Attitude towards physics										
Pearson correlations	0.191	0.227*	1	-0.322**	0.412**	-0.038	0.213*	-0.124	0.184	-0.101
Sig. (2-tailed)	0.068	0.030		0.002	0.000	0.717	0.042	0.239	0.079	0.339
Confidence										
Pearson correlations	-0.322**	-0.213*	-0.332**	1	-0.166	-0.097	-0.165	0.243*	-0.264*	0.244*
Sig. (2-tailed)	0.002	0.042	0.001		0.113	0.360	0.116	0.020	0.011	0.019
Source of motivation										
Pearson correlations	0.172	0.316**	0.412**	-0.166	1	-0.049	0.092	-.0.008	0.213*	0.086
Sig. (2-tailed)	0.100	0.002	0.000	0.113		0.642	0.385	0.939	0.042	0.417
Prospective level of difficulty										
Pearson correlations	-0.026	0.136	-0.038	-0.097	-0.049	1	0.054	0.055	0.152	-0.010
Sig. (2-tailed)	0.807	0.197	0.717	0.360	0.642		0.610	0.603	0.148	0.926

Table K.6 Pearson correlations of the prospective outcome emotion variables part 1

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'.

	Gender	Pre-test	Attitude towards physics	Confidence	Source of motivation	Prospective level of difficulty	Attitude towards effort	Emotion	Value	Control
Attitude towards effort										
Pearson correlations	0.255*	0.270**	0.213*	-0.165	0.092	0.054	1	-0.015	0.037	0.003
Sig. (2-tailed)	0.014	0.009	0.042	0.116	0.385	0.610		0.885	0.727	0.981
Emotion										
Pearson correlations	-0.195	0.114	-0.124	0.243*	-.0.008	0.055	-0.015	1	0.224*	0.928**
Sig. (2-tailed)	0.063	0.278	0.239	0.020	0.939	0.603	0.885		0.032	0.000
Value										
Pearson correlations	-0.006	0.408**	0.184	-0.264*	0.213*	0.152	0.037	0.224*	1	0.256*
Sig. (2-tailed)	0.957	0.000	0.079	0.011	0.042	0.148	0.727	0.032		0.014
Control										
Pearson correlations	-0.153	0.121	-0.101	0.244*	0.086	-0.010	0.003	0.928**	0.256*	1
Sig. (2-tailed)	0.145	0.252	0.339	0.019	0.417	0.926	0.981	0.000	0.014	

Table K.7 Pearson correlations of the prospective outcome emotion variables part 2

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'.

K.3 Results of BLR/MLR for activity emotions

On this section are presented the results of Binary Logistic Regression for *control* and Multinomial Logistic Regression for *value for 708 cases*, when the regressors or independent variables are discretised into two and three categories. The method employed to select the random variables was the 'Forward Conditional' or the 'Stepwise' procedure in SPSS.

Value of the *activity emotions*							
Observed	Predicted						
	Negative	None	Positive	Specificity	Sensitivity	Precision	Overall model accuracy
Negative	270	3	51	0.341	0.833	0.516	0.528
None	91	5	26	0.993	0.041	0.556	
Positive	162	1	99	0.827	0.378	0.563	

Table K.8 Confusion matrix of *value* with regressors divided into two categories[2]

Control of the *activity* emotions						
Observed	Predicted					
	High	Low	Specificity	Sensitivity	Precision	Overall model accuracy
High	277	121	0.452	0.696	0.620	0.589
Low	170	140	0.696	0.452	0.530	

Table K.9 Confusion matrix of *control* with regressors divided into two categories[3]

Value of the *activity emotions*							
Observed	Predicted						
	Negative	None	Positive	Specificity	Sensitivity	Precision	Overall model accuracy
Negative	239	2	83	0.555	0.738	0.583	0.595
None	90	2	30	0.995	0.016	0.400	
Positive	81	1	180	0.747	0.687	0.614	

Table K.10 Confusion matrix of *value* (regressors in two categories, control & value t-1)

[2] The regressors chosen are outcome, times asked help and focus coarse value

[3] The regressors chosen are outcome, attempts alone, average quality tutoring feedback, focus coarse value

Control of the *activity* emotions						
Observed	Predicted					
	High	Low	Specificity	Sensitivity	Precision	Overall model accuracy
High	308	90	0.613	0.774	0.719	0.703
Low	120	190	0.774	0.613	0.678	

Table K.11 Confusion matrix of control (variables in two categories, *control* & *value t-1)*

Value of the *activity emotions*							
Observed	Predicted						
	Negative	None	Positive	Specificity	Sensitivity	Precision	Overall model accuracy
Negative	237	0	87	0.466	0.731	0.536	0.537
None	86	0	36	1.000	0.000	Undefined	
Positive	119	0	143	0.724	0.546	0.538	

Table K.12 Confusion matrix of *value* with regressors divided into three categories[4]

Control of the *activity* emotions						
Observed	Predicted					
	High	Low	Specificity	Sensitivity	Precision	Overall model accuracy
High	307	91	0.361	0.771	0.608	0.592
Low	198	112	0.771	0.361	0.552	

Table K.13 Confusion matrix of *control* with regressors divided into three categories[5]

Value of the *activity emotions*							
Observed	Predicted						
	Negative	None	Positive	Specificity	Sensitivity	Precision	Overall model accuracy
Negative	249	2	73	0.568	0.769	0.600	0.617
None	91	2	29	0.995	0.016	0.400	
Positive	75	1	186	0.771	0.710	0.646	

Table K.14 Confusion matrix of *value* (regressors in two categories, *control & value t-1)*

[4] The regressors chosen are overall attempts, times asked help and average quality tutoring feedback

[5] The regressors chosen are outcome, overall attempts, average quality tutoring feedback, focus coarse value

Control of the *activity* emotions						
Observed	Predicted					
	High	Low	Specificity	Sensitivity	Precision	Overall model accuracy
High	270	128	0.529	0.678	0.649	0.6129
Low	146	164	0.678	0.529	0.562	

Table K.15 Confusion matrix of *control* (variables in 3 categories*, control & value t-1)*

K.4 CPTs corresponding to activity emotions network

Type of outcome	
InProgress	0.986
BlackOut	0.002
PositiveAcceleration	0.012

Average quality tutoring feedback	
'(1.765-inf)'	0.848
'(-inf-1.765]'	0.152

Table K.16 CPTs of *type of outcome & average quality of tutoring feedback*

Overall attempts	*Attitude towards physics*	
	Positive	Negative
'(-inf-1.5]'	0.800	0.250
'(1.5-inf)'	0.200	0.750

Time to achieve learn-ing goals	*Interval of interaction*	
	'(559-inf)'	'(-inf-559]'
'(-inf-1242.301724]'	0.850	1.000
'(1242.301724-inf)'	0.149	0.000

Table K.17 CPTs of *total attempts & time to achieve learning goals*

Estimated Independence		'(-inf-0.5]'		'(0.5-inf)'	
Interval of interaction		'(559-inf)'	'(-inf-559]'	'(559-inf)'	'(-inf-559]'
Overall Attempts	'(-inf-1.5]'	0.097	0.850	0.135	0.582
	'(1.5-inf)'	0.903	0.149	0.865	0.418

Table K.18 CPT corresponding to *total attempts*

Outcome	*Type of outcome*		
	InProgress	BlackOut	PositiveAcceleration
'(9.285-inf)'	1.000	1.000	0.000
'(-inf-9.285]'	0.000	0.000	1.000

Table K.19 CPT corresponding to *outcome (progress) and value t-1*

Estimated independence	
'(-inf-0.5]'	0.521
'(0.5-inf)'	0.479

Table K.20 CPT corresponding to *estimated independence*

Value t-1	
Negative	0.575
Positive	0.321
None	0.104

Table K.21 CPT corresponding to *value t-1*

Interval of interaction	*Average quality tutoring feedback*	
	'(1.765-inf)'	'(-inf-1.765]'
'(559-inf)'	0.381	0.961
'(-inf-559]'	0.619	0.039

Estimated independence		'(-inf-0.5]'		'(0.5-inf)'	
Overall attempts		'(-inf-1.5]'	'(1.5-inf)'	'(-inf-1.5]'	'(1.5-inf)'
Attempts alone	'(-inf-1.5]'	1.000	0.523	1.000	0.000
	'(1.5-inf)'	0.000	0.477	0.000	1.000

Table K.22 CPTs corresponding to *interval of interaction and attempts alone*

Control		Low			High		
Value		Negative	Positive	None	Negative	Positive	None
Emotion	Frustration	1.000	0.250	0.000	0.000	0.000	0.250
	Enjoyment	0.000	0.250	0.000	0.000	1.000	0.250
	Anger	0.000	0.250	0.000	1.000	0.000	0.250
	Boredom	0.000	0.250	1.000	0.000	0.000	0.250

Table K.23 CPT corresponding to *activity achievement emotions*

Estimated Independence		'(-inf-0.5]'						'(0.5-inf)'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control t-1	Low	0.536	0.000	1.000	0.318	0.000	1.000	0.326	0.000	1.000	0.300	0.000	1.000
	Self	0.000	0.016	0.000	0.141	0.000	0.000	0.256	0.048	0.000	0.100	0.070	0.000
	High	0.319	0.703	0.000	0.318	0.833	0.000	0.163	0.762	0.000	0.267	0.789	0.000
	Other	0.072	0.031	0.000	0.129	0.000	0.000	0.186	0.095	0.000	0.222	0.070	0.000
	Irrelevant	0.043	0.000	0.000	0.094	0.167	0.000	0.069	0.048	0.000	0.111	0.070	0.000
	Medium	0.029	0.250	0.000	0.000	0.000	0.000	0.000	0.048	0.000	0.000	0.000	0.000

Table K.24 CPT corresponding to *control t-1*

Average quality tutoring feedback		'(1.765-inf)'											
Estimated independence		'(-inf-0.5]'											
Interval of interaction		'(559-inf)'						'(-inf-559)'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Times asked help	'(0.5-inf)'	1.000	1.000	0.500	1.000	1.000	1.000	0.508	0.066	0.278	1.000	1.000	1.000
	'(-inf-0.5]'	0.000	0.000	0.500	0.000	0.000	0.000	0.492	0.934	0.722	0.000	0.000	0.000

Table K.25 CPT corresponding to *times asked help* part1

Average quality tutoring feedback		'(1.765-inf)'											
Estimated independence		'(0.5-inf)'											
Interval of interaction		'(559-inf)'						'(-inf-559)'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Times asked help	'(0.5-inf)'	0.667	1.000	0.500	0.383	0.094	0.333	0.161	0.222	0.125	0.158	0.333	0.000
	'(-inf-0.5]'	0.333	0.000	0.500	0.617	0.906	0.667	0.839	0.778	0.875	0.842	0.667	1.000

Table K.26 CPT corresponding to *times asked help* part 2

Average quality tutoring feedback		'(-inf-1.765]'											
Estimated independence		'(-inf-0.5]'											
Interval of interaction		'(559-inf)'						'(-inf-559)'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Times asked help	'(0.5-inf)'	0.500	0.500	0.500	1.000	1.000	1.000	0.500	0.500	0.500	1.000	0.500	0.500
	'(-inf-0.5]'	0.500	0.500	0.500	0.000	0.000	0.000	0.500	0.500	0.500	0.000	0.500	0.500

Table K.27 CPT corresponding to *times asked help* part 3

Average quality tutoring feedback		'(-inf-1.765]'											
Estimated independence		'(0.5-inf)'											
Interval of interaction		'(559-inf)'						'(-inf-559)'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Times asked help	'(0.5-inf)'	1.000	0.500	1.000	1.000	1.000	1.000	0.500	0.500	0.500	0.500	1.000	0.500
	'(-inf-0.5]'	0.000	0.500	0.000	0.000	0.000	0.000	0.500	0.500	0.500	0.500	0.000	0.500

Table K.28 CPT corresponding to times asked help part 4

Control t-1		Low											
Times asked help		'(0.5-inf)'						'(-inf-0.5]'					
Outcome		'(9.285-inf)'			'(-inf-9.285]'			'(9.285-inf)'			'(-inf-9.285]'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Value	Negative	0.800	0.333	0.550	1.000	0.333	0.000	0.564	0.333	0.452	0.333	0.333	0.333
	Positive	0.046	0.333	0.100	0.000	0.333	0.000	0.051	0.333	0.258	0.333	0.333	0.333
	None	0.154	0.333	0.350	0.000	0.333	1.000	0.385	0.333	0.290	0.333	0.333	0.333

Table K.29 CPT corresponding to *value* part 1

Control t-1		Self											
Times asked help		'(0.5-inf)'						'(-inf-0.5]'					
Outcome		'(9.285-inf)'			'(-inf-9.285]'			'(9.285-inf)'			'(-inf-9.285]'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Value	Negative	0.571	0.000	0.333	0.000	0.333	0.333	0.444	0.200	0.333	0.333	0.333	0.333
	Positive	0.429	0.000	0.333	0.000	0.333	0.333	0.333	0.800	0.333	0.333	0.333	0.333
	None	0.000	1.000	0.333	1.000	0.333	0.333	0.222	0.000	0.333	0.333	0.333	0.333

Table K.30 CPT corresponding to value part 2

Control t-1		High											
Times asked help		'(0.5-inf)'						'(-inf-0.5]'					
Outcome		'(9.285-inf)'			'(-inf-9.285]'			'(9.285-inf)'			'(-inf-9.285]'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Value	Negative	0.711	0.209	0.333	0.333	0.333	0.333	0.412	0.269	0.333	0.000	0.333	0.333
	Positive	0.111	0.628	0.333	0.333	0.333	0.333	0.324	0.679	0.333	0.000	0.333	0.333
	None	0.178	0.163	0.333	0.333	0.333	0.333	0.265	0.051	0.333	1.000	0.333	0.333

Table K.31 CPT corresponding to *value* part 3

Control t-1		Other											
Times asked help		'(0.5-inf)'						'(-inf-0.5]'					
Outcome		'(9.285-inf)'			'(-inf-9.285]'			'(9.285-inf)'			'(-inf-9.285]'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Value	Negative	0.542	0.333	0.333	0.333	0.333	0.333	0.400	0.250	0.333	0.333	0.333	0.333
	Positive	0.333	0.333	0.333	0.333	0.333	0.333	0.600	0.625	0.333	0.333	0.333	0.333
	None	0.125	0.333	0.333	0.333	0.333	0.333	0.000	0.125	0.333	0.333	0.333	0.333

Table K.32 CPT corresponding to *value* part 4

Control t-1		Irrelevant											
Times asked help		'(0.5-inf)'						'(-inf-0.5]'					
Outcome		'(9.285-inf)'			'(-inf-9.285]'			'(9.285-inf)'			'(-inf-9.285]'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Value	Negative	0.353	0.250	0.333	0.333	0.333	0.333	0.286	0.250	0.333	0.333	0.333	0.333
	Positive	0.353	0.250	0.333	0.333	0.333	0.333	0.286	0.500	0.333	0.333	0.333	0.333
	None	0.294	0.500	0.333	0.333	0.333	0.333	0.428	0.250	0.333	0.333	0.333	0.333

Table K.33 CPT corresponding to *value* part 5

Control t-1		Medium											
Times asked help		'(0.5-inf)'						'(-inf-0.5]'					
Outcome		'(9.285-inf)'			'(-inf-9.285]'			'(9.285-inf)'			'(-inf-9.285]'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Value	Negative	0.333	0.333	0.333	0.333	0.333	0.333	0.500	0.125	0.333	0.333	0.000	0.333
	Positive	0.333	0.333	0.333	0.333	0.333	0.333	0.500	0.625	0.333	0.333	1.000	0.333
	None	0.333	0.333	0.333	0.333	0.333	0.333	0.000	0.250	0.333	0.333	0.000	0.333

Table K.34 CPT corresponding to *value* part 6

Average quality tutoring feedback		'(1.765-inf)'											
Control t-1		Low											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.841	0.500	0.577	0.441	0.500	0.895	1.000	0.500	0.500	0.500	0.500	1.000
	High	0.159	0.500	0.423	0.559	0.500	0.105	0.000	0.500	0.500	0.500	0.500	0.000

Table K.35 CPT corresponding to *control* part 1

Average quality tutoring feedback		'(1.765-inf)'											
Control t-1		Self											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.545	0.000	0.500	0.231	0.500	0.500	0.500	0.500	0.500	1.000	0.500	0.500
	High	0.454	1.000	0.500	0.769	0.500	0.500	0.500	0.500	0.500	0.000	0.500	0.500

Table K.36 CPT corresponding to *control* part 2

Average quality tutoring feedback		'(1.765-inf)'											
Control t-1		High											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.358	0.197	0.500	0.342	0.163	0.500	1.000	0.500	0.500	0.500	0.500	0.500
	High	0.643	0.803	0.500	0.658	0.837	0.500	0.000	0.500	0.500	0.500	0.500	0.500

Table K.37 CPT corresponding to *control* part 3

Average quality tutoring feedback		'(1.765-inf)'											
Control t-1		Other											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.308	0.250	0.500	0.091	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	High	0.692	0.750	0.500	0.909	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500

Table K.38 CPT corresponding to *control* part 4

Average quality tutoring feedback		'(1.765-inf)'											
Control t-1		Irrelevant											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.833	1.000	0.500	0.533	0.600	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	High	0.167	0.000	0.500	0.467	0.400	0.500	0.500	0.500	0.500	0.500	0.500	0.500

Table K.39 CPT corresponding to *control* part 5

Average quality tutoring feedback		'(1.765-inf)'											
Control t-1		Medium											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.500	0.313	0.500	0.500	0.500	0.500	0.500	0.000	0.500	0.500	0.500	0.500
	High	0.500	0.687	0.500	0.500	0.500	0.500	0.500	1.000	0.500	0.500	0.500	0.500

Table K.40 CPT corresponding to *control* part 6

Average quality tutoring feedback		'(-inf-1.765]'											
Control t-1		Low											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	1.000	0.500	0.000	0.900	0.500	0.000	0.500	0.500	0.500	0.500	0.500	0.500
	High	0.000	0.500	1.000	0.100	0.500	0.100	0.500	0.500	0.500	0.500	0.500	0.500

Table K.41 CPT corresponding to *control* part 7

Average quality tutoring feedback		'(-inf-1.765]'											
Control t-1		Self											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.500	0.500	0.500	0.833	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	High	0.500	0.500	0.500	0.167	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500

Table K.42 CPT corresponding to *control* part 8

Average quality tutoring feedback		'(-inf-1.765]'											
Control t-1		High											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.500	0.500	0.500	0.538	0.364	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	High	0.500	0.500	0.500	0.462	0.636	0.500	0.500	0.500	0.500	0.500	0.500	0.500

Table K.43 CPT corresponding to *control* part 9

Average quality tutoring feedback		'(-inf-1.765]'											
Control t-1		Other											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.500	0.500	0.500	0.667	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	High	0.500	0.500	0.500	0.333	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500

Table K.44 CPT corresponding to *control* part 10

Average quality tutoring feedback		'(-inf-1.765]'											
Control t-1		Irrelevant											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.500	0.500	0.500	0.667	0.000	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	High	0.500	0.500	0.500	0.333	1.000	0.500	0.500	0.500	0.500	0.500	0.500	0.500

Table K.45 CPT corresponding to *control* part 11

Average quality tutoring feedback		'(-inf-1.765]'											
Control t-1		Medium											
Outcome		'(9.285-inf)'						'(-inf-9.285]'					
Overall attempts		'(-inf-1.5]'			'(1.5-inf)'			'(-inf-1.5]'			'(1.5-inf)'		
Value t-1		Negative	Positive	None	Negative	Positive	None	Negative	Positive	None	Negative	Positive	None
Control	Low	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	High	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500

Table K.46 CPT corresponding to *control* part 12

K.5 Pearson correlations of activity emotions variables

The Pearson correlations for the variables involved in the creation of the model corresponding to the prospective outcome emotions are presented for 499 cases.

	Outcome	Type of outcome	Times asked help	Attempts alone	Estimated independ-ence	Overall attempts	Average quality tu-toring feedback	Interval of inter-action	Focus coarse value	Time to achieve learning goals	Emotion	Control	Value	Control t-1	Value t-1
Outcome															
Pearson corre-lations	1	-0.925**	-0.039	0.022	0.032	0.010	-0.047	0.104*	0.045	0.030	0.017	-0.086	-0.013	-0.083	0.030
Sig. (2-tailed)		1.4E-211	0.382	0.628	0.474	0.824	0.297	0.021	0.320	0.499	0.711	0.056	0.766	0.064	0.508
Type of out-come															
Pearson corre-lations	-0.925**	1	0.019	-0.040	-0.013	-0.025	0.083	-0.114*	-0.057	-0.023	-0.021	0.094*	-0.009	0.052	-0.014
Sig. (2-tailed)	1.4E-211		0.674	0.378	0.765	0.581	0.063	0.011	0.203	0.603	0.644	0.035	0.846	0.248	0.763

Table K.47 Pearson correlations of activity emotion variables part 1

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'..

	Outcome	Type of outcome	Times asked help	Attempts alone	Estimated independ-ence	Overall attempts	Average quality tu-toring feedback	Interval of inter-action	Focus coarse value	Time to achieve learning goals	Emotion	Control	Value	Control t-1	Value t-1
Times asked help															
Pearson cor-relations	-0.039	0.019	1	0.039	-0.256**	0.274**	-0.433**	0.350**	-0.053	0.061	-0.046	0.096*	-0.223**	0.048	-0.285**
Sig. (2-tailed)	0.382	0.674		0.380	6.8E-9	5.03E-10	2.94E-24	7.8E-16	0.240	0.174	0.301	0.031	5.1E-7	0.289	8.3E-11
Attempts alone															
Pearson cor-relations	0.022	-0.040	0.039	1	0.486**	0.795**	-0.082	0.512**	0.081	0.268**	-0.164**	-0.098*	-0.003	0.017	-0.055
Sig. (2-tailed)	0.628	0.378	0.380		5.6E-31	6.7E-110	0.068	9.6E-35	0.071	1.2E-9	2.4E-4	0.028	0.941	0.704	0.224
Estimated independ-ence															
Pearson cor-relations	0.032	-0.013	-0.256**	0.486**	1	0.264**	0.083	0.233**	-0.079	0.114*	-0.098*	-0.047	0.076	0.089*	0.026
Sig. (2-tailed)	0.474	0.765	6.8E-9	5.6E-31		2.2E-9	0.065	1.5E-7	0.076	0.011	0.029	0.296	0.091	0.046	0.568
Overall at-tempts															
Pearson cor-relations	0.010	-0.025	0.274**	0.795**	0.264**	1	-0.275**	0.633**	0.052	0.251**	-0.143**	-0.029	-0.039	0.013	-0.136**
Sig. (2-tailed)	0.824	0.581	5.03E-10	6.7E-110	2.2E-9		4.01E-10	3.8E-57	0.245	1.33E-8	0.001	0.521	0.390	0.767	0.002

Table K.48 Pearson correlations of activity emotion variables part 2

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'.

	Outcome	Type of outcome	Times asked help	Attempts alone	Estimated independ-ence	Overall attempts	Average quality tu-toring feedback	Interval of inter-action	Focus coarse value	Time to achieve learning goals	Emotion	Control	Value	Control t-1	Value t-1
Average quality tutor-ing feedback															
Pearson cor-relations	-0.047	0.083	-0.433**	-0.082	0.083	-0.275**	1	-0.418**	0.063	-0.211**	-0.082	-0.157**	0.098*	-0.044	0.151**
Sig. (2-tailed)	0.297	0.063	2.94E-24	0.068	0.065	4.01E-10		1.8E-22	0.158	1.9E-6	0.067	4.15E-4	0.028	0.321	0.001
Interval of in-teraction															
Pearson cor-relations	0.104*	-0.114*	0.350**	0.512**	0.233**	0.633**	-0.418**	1	-0.066	0.292**	-0.074	0.020	-0.088*	-0.009	-0.161**
Sig. (2-tailed)	0.021	0.011	7.8E-16	9.6E-35	1.5E-7	3.8E-57	1.8E-22		0.141	2.8E-11	0.098	0.662	0.048	0.843	3.2E-4
Focus coarse value															
Pearson cor-relation	0.045	-0.057	-0.053	0.081	-0.079	0.052	0.063	-0.066	1	-0.090*	-0.138**	-0.003	-0.131**	-0.085	-0.084*
Sig, (2-tailed)	0.320	0.203	0.240	0.071	0.076	0.245	0.158	0.141		0.044	0.002	0.954	0.003	0.057	0.061
Time to achieve learn-ing goals															
Pearson cor-relation	0.030	-0.023	0.061	0.268**	0.114*	0.251**	-0.211**	0.292**	-0.090*	1	-0.075	-0.026	-0.063	-0.025	0.034
Sig, (2-tailed)	0.499	0.603	0.174	1.2E-9	0.011	1.33E-8	1.9E-6	2.8E-11	0.044		0.096	0.564	0.158	0.578	0.453

Table K.49 Pearson correlations of activity emotion variables part 3

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'.

	Outcome	Type of outcome	Times asked help	Attempts alone	Estimated independ-ence	Overall attempts	Average quality tu-toring feedback	Interval of inter-action	Focus coarse value	Time to achieve learning goals	Emotion	Control	Value	Control t-1	Value t-1
Emotion Pearson cor-relation	0.017	-0.021	-0.046	-0.164**	-0.098*	-0.143**	-0.082	-0.074	-0.138**	-0.075	1	0.409**	0.117**	0.125**	0.056
Sig. (2 tailed)	0.711	0.644	0.301	2.4E-4	0.029	0.001	0.067	0.098	0.002	0.096		1.4E-21	0.009	0.005	0.216
Control Pearson cor-relation	-0.086	0.094*	0.096*	-0.098*	-0.047	-0.029	-0.157**	0.020	-0.003	-0.026	0.409**	1	-0.469**	0.121**	-0,253**
Sig. (2 tailed)	0.056	0.035	0.031	0.028	0.296	0.521	4.15E-4	0.662	0.954	0.564	1.4E-21		1.29E-28	0.007	9.8E-9
Value Pearson cor-relation	-0.013	-0.009	-0.223**	-0.003	0.076	-0.039	0.098*	-0.088*	-0.131**	-0.063	0.117**	-0.469**	1	-0.072	0.387**
Sig. (2-tailed)	0.766	0.846	5.1E-7	0.941	0.091	0.390	0.028	0.048	0.003	0.158	0.009	1.29E-28		0.108	2.82E-19
Control t-1 Pearson cor-relation	-0.083	0.052	0.048	0.017	0.089*	0.013	-0.044	-0.009	-0.085	-0.025	0.125**	0.121**	-0.072	1	-0.339**
Sig. (2-tailed)	0.064	0.248	0.289	0.704	0.046	0.767	0.321	0.843	0.057	0.578	0.005	0.007	0.108		7.6E-15
Value t-1 Pearson cor-relation	0.030	-0.014	-0.285**	-0.055	0.026	-0.136**	0.151**	-0.161**	-0.084*	0.034	0.056	-0,253**	0.387**	-0.339**	1
Sig. (2-tailed)	0.508	0.763	8.3E-11	0.224	0.568	0.002	0.001	3.2E-4	0.061	0.453	0.216	9.8E-9	2.82E-19	7.6E-15	

Table K.50 Pearson correlations of activity emotion variables part 4

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'.

K.6 Pearson correlations of retrospective outcome emotions variables

The Pearson correlations for the variables involved in the creation of the model corresponding to the prospective outcome emotions are presented for 259 cases.

	Outcome	Type of outcome	Times asked help	Attempts alone	Estimated independ-ence	Overall at-tempts	Average quality tutoring feedback	Interval of interaction	Focus coarse value	Time to achieve learning goals	Publishing outcome	Emotion	Control	Value	Control t-1	Value t-1
Outcome Pearson correlations	1	0.308**	0.034	0.231**	0.199**	0.225**	-0.029	0.336**	-0.143*	-0.557**	-0.416**	-0.072	0.048	0.311**	-0.048	0.181
Sig. (2-tailed)		4.26E-7	0.584	1.78E-4	0.001	2.59E-4	0.646	2.93E-8	0.021	1.61E-22	3.06E-12	0.248	0.442	3.22E-7	0.443	0.193
Type of outcome Pearson correlations	0.308**	1	-0.033	-0.081	-0.042	-0.075	0.035	0.059	0.079	-0.386**	-0.429**	0.027	0.079	0.223**	-0.076	0.080
Sig. (2-tailed)	4.26E-7		0.602	0.194	0.504	0.227	0.570	0.340	0.203	1.25E-10	4.75E-13	0.668	0.205	2.94E-4	0.222	0.197
Times asked help Pearson correlations	0.034	-0.033	1	0.167**	-0.352**	0.368**	-0.130*	0.371**	-0.159*	-0.025	0.020	-0.152*	0.045	-0.064	0.123*	-0.223**
Sig. (2-tailed)	0.584	0.602		0.007	5.45E-9	9.84E-10	0.037	7.36E-10	0.010	0.690	0.752	0.014	0.467	0.307	0.048	2.89E-4

Table K.51 Pearson correlations of retrospective outcome emotion variables part 1

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'.

	Outcome	Type of outcome	Times asked help	Attempts alone	Estimated independ-ence	Overall at-tempts	Average quality tu-toring feed-back	Interval of interaction	Focus coarse value	Time to achieve learning goals	Publishing outcome	Emotion	Control	Value	Control t-1	Value t-1
Attempts alone																
Pearson correlations	0.231**	-0.081	0.167**	1	0.461**	0.792**	-0.044	0.519**	-0.169**	-0.091	-0.049	-0.163**	0.029	0.038	0.035	-0.060
Sig. (2-tailed)	1.78E-4	0.194	0.007		4.58E-15	4.83E-57	0.480	2.98E-19	0.007	0.145	0.434	0.009	0.644	0.547	0.573	0.336
Estimated independ-ence																
Pearson correlations	0.199**	-0.042	-0.352**	0.461**	1	0.261**	-0.018	0.176**	-0.059	-0.118	-0.101	-0.028	0.019	0.136*	0.040	-0.021
Sig. (2-tailed)	0.001	0.504	5.45E-9	4.58E-15		2.07E-5	0.775	0.005	0.341	0.057	0.105	0.657	0.766	0.029	0.525	0.735
Overall at-tempts																
Pearson correlations	0.225**	-0.075	0.368**	0.792**	0.261**	1	-0.137*	0.608**	-0.151*	-0.092	-0.092	0.208**	0.046	0.092	0.074	-0.083
Sig. (2-tailed)	2.59E-4	0.227	9.84E-10	4.83E-57	2.07E-5		0.028	1.26E-27	0.015	0.138	0.138	0.001	0.461	0.138	0.235	0.185
Average quality tu-toring feed-back																
Pearson correlations	-0.029	0.035	-0.130*	-0.044	-0.018	-0.137*	1	-0.078	-0.064	0.024	0.003	0.100	-0.093	-0.088	-0.103	-0.011
Sig. (2-tailed)	0.646	0.570	0.037	0.480	0.775	0.028		0.209	0.303	0.702	0.957	0.109	0.134	0.157	0.097	0.865

Table K.52 Pearson correlations of retrospective outcome emotion variables part 2

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'

	Outcome	Type of outcome	Times asked help	Attempts alone	Estimated independ-ence	Overall at-tempts	Average quality tutoring feedback	Interval of interaction	Focus coarse value	Time to achieve learning goals	Publishing outcome	Emotion	Control	Value	Control t-1	Value t-1
Interval of interaction																
Pearson correlations	0.336**	0.059	0.371**	0.519**	0.176**	0.608**	-0.078	1	-0.251**	-0.090	-0.081	-0.133*	0.047	0.073	0.009	-0.026
Sig. (2-tailed)	2.93E-8	0.340	7.36E-10	2.98E-19	0.005	1.26E-27	0.209		4.40E-5	0.148	0.196	0.033	0.450	0.245	0.889	0.676
Focus coarse value																
Pearson correlations	-0.143*	0.079	-0.159*	-0.169**	-0.059	-0.151*	-0.064	-0.251**	1	0.068	-0.058	0.090	0.058	-0.015	-0.013	-0.009
Sig. (2-tailed)	0.021	0.203	0.010	0.007	0.341	0.015	0.303	4.40E-5		0.276	0.353	0.148	0.356	0.807	0.835	0.882
Time to achieve learning goals																
Pearson correlations	-0.557**	-0.386**	-0.025	-0.091	-0.118	-0.092	0.024	-0.090	0.068	1	0.457**	0.117	-0.098	-0.282**	0.084	-0.049
Sig. (2-tailed)	1.61E-22	1.25E-10	0.690	0.145	0.057	0.138	0.702	0.148	0.276		8.86E-15	0.061	0.117	3.99E-6	0.178	0.428
Publishing outcome																
Pearson correlations	-0.416**	-0.429**	0.020	-0.049	-0.101	-0.092	0.003	-0.081	-0.058	0.457**	1	-0.090	-0.177**	-0.508**	0.064	-0.065
Sig. (2-tailed)	3.06E-12	4.75E-13	0.752	0.434	0.105	0.138	0.957	0.196	0.353	8.86E-15		0.148	0.004	1.95E-18	0.308	0.297

Table K.53 Pearson correlations of retrospective outcome emotion variables part 3

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'

	Outcome	Type of outcome	Times asked help	Attempts alone	Estimated independ-ence	Overall at-tempts	Average quality tutoring feedback	Interval of interaction	Focus coarse value	Time to achieve learning goals	Publishing outcome	Emotion	Control	Value	Control t-1	Value t-1
Emotion Pearson correlations	-0.072	0.027	-0.152*	-0.163**	-0.028	0.208**	0.100	-0.133*	0.090	0.117	-0.090	1	0.163**	0.047	-0.071	0.145*
Sig. (2-tailed)	0.248	0.668	0.014	0.009	0.657	0.001	0.109	0.033	0.148	0.061	0.148		0.008	0.447	0.252	0.020
Control Pearson correlations	0.048	0.079	0.045	0.029	0.019	0.046	-0.093	0.047	0.058	-0.098	-0.177**	0.163**	1	0.086	0.077	0.038
Sig. (2-tailed)	0.442	0.205	0.467	0.644	0.766	0.461	0.134	0.450	0.356	0.117	0.004	0.008		0.169	0.219	0.546
Value Pearson correlations	0.311**	0.223**	-0.064	0.038	0.136*	0.092	-0.088	0.073	-0.015	-0.282**	-0.508**	0.047	0.086	1	-0.133*	0.221**
Sig. (2-tailed)	3.22E-7	2.94E-4	0.307	0.547	0.029	0.138	0.157	0.245	0.807	3.99E-6	1.95E-18	0.447	0.169		0.032	3.32E-4
Control t-1 Pearson correlations	-0.048	-0.076	0.123*	0.035	0.040	0.074	-0.103	0.009	-0.013	0.084	0.064	-0.071	0.077	-0.133*	1	-0.421**
Sig. (2-tailed)	0.443	0.222	0.048	0.573	0.525	0.235	0.097	0.889	0.835	0.178	0.308	0.252	0.219	0.032		1.54E-12
Value t-1 Pearson correlations	0.181	0.080	-0.223**	-0.060	-0.021	-0.083	-0.011	-0.026	-0.009	-0.049	-0.065	0.145*	0.038	0.221**	-0.421**	1
Sig. (2-tailed)	0.193	0.197	2.89E-4	0.336	0.735	0.185	0.865	0.676	0.882	0.428	0.297	0.020	0.546	3.32E-4	1.54E-12	

Table K.54 Pearson correlations of retrospective outcome emotion variables part 4

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'

K.7 Results of BLR/MLR for retrospective outcome emotions

On this section are presented the results of Binary Logistic Regression for *value* and Multinomial Logistic Regression for *control for 259 cases*, when the regressors or independent variables are discretised into two and three categories. The method employed to select the variables was the 'Forward Conditional' or the 'Stepwise' procedure in SPSS.

Control of the *retrospective outcome emotions*							
Observed	Predicted						
	Irrelevant	Other	Self	Specificity	Sensitivity	Precision	Overall model accuracy
Irrelevant	0	61	8	0.000	1.000	Undefined	0.162
Other	0	111	7	0.041	0.941	0.143	
Self	0	52	20	0.979	0.278	0.571	

Table K.55 Confusion matrix of *control* with regressors divided into two categories [6]

Value of retrospective outcome emotions						
Observed	Predicted					
	Negative	Positive	Specificity	Sensitivity	Precision	Overall model accuracy
Negative	184	7	0.485	0.963	0.840	0.838
Positive	35	33	0.963	0.485	0.825	

Table K.56 Confusion matrix of *value* with regressors divided into two categories[7]

Control of the *retrospective outcome emotions*							
Observed	Predicted						
	Irrelevant	Other	Self	Specificity	Sensitivity	Precision	Overall model accuracy
Irrelevant	20	36	13	0.932	0.289	0.606	0.579
Other	9	98	11	0.489	0.831	0.576	
Self	4	36	32	0.872	0.444	0.571	

Table K.57 Confusion matrix of *control* (regressors in two categories & *control t-1)*

[6] The regressors chosen are publishing outcome, average quality tutoring feedback and time to achieving learning goals

[7] The regressors chosen are publishing outcome and type of outcome

Value of the retrospective outcome emotions						
Observed	Predicted					
	Negative	Positive	Specificity	Sensitivity	Precision	Overall model accuracy
Negative	185	6	0.426	0.969	0.823	0.826
Positive	39	29	0.969	0.426	0.829	

Table K.58 Confusion matrix of *value* (regressors in two categories & *control t-1)*

Control of the *retrospective outcome emotions*							
Observed	Predicted						
	Irrelevant	Other	Self	Specificity	Sensitivity	Precision	Overall model accuracy
Irrelevant	0	69	0	1.000	0.000	Undefined	0.456
Other	0	118	0	0.000	1.000	0.456	
Self	0	72	0	1.000	0.000	Undefined	

Table K.59 Confusion matrix of *control* with regressors divided into three categories [8]

Value of the retrospective outcome emotions						
Observed	Predicted					
	Negative	Positive	Specificity	Sensitivity	Precision	Overall model accuracy
Negative	185	6	0.426	0.969	0.823	0.838
Positive	39	29	0.969	0.426	0.829	

Table K.60 Confusion matrix of *value* with regressors divided into three categories[9]

Control of the *retrospective outcome emotions*							
Observed	Predicted						
	Irrelevant	Other	Self	Specificity	Sensitivity	Precision	Overall model accuracy
Irrelevant	32	24	13	0.832	0.464	0.5	0.579
Other	21	86	11	0.624	0.728	0.619	
Self	11	29	32	0.872	0.444	0.571	

Table K.61 Confusion matrix of *control* (regressors in three categories & *control t-1*)

[8] The regressors chosen are times asked help

[9] The regressors chosen are publishing outcome and type of outcome

Value of retrospective outcome emotions						
Observed	Predicted					
	Negative	Positive	Specificity	Sensitivity	Precision	Overall model accuracy
Negative	185	6	0.426	0.969	0.826	0.826
Positive	39	29	0.969	0.426	0.826	

Table K.62 Confusion matrix of *value* (regressors in three categories & *control t-1*)

K.8 CPTs corresponding to retrospective outcome emotions network

Time to achieve learning goals	
'(1981.664-inf)'	0.737
'(-inf-1981.664]'	0.263

Interval of interaction	
'(773-inf)'	0.498
'(-inf-773]'	0.502

Table K.63 CPTs of *time to achieve learning goals* and *interval of interaction*

Outcome	*Type of outcome*						
	PositiveAcceleration	BlackOut	Success	Distance-TooFar	Distance-TooShort	Time-Out	Re-Started
'(-inf-55]'	1.000	1.000	0.000	0.000	0.000	0.947	1.000
'(55-inf)'	0.000	0.000	1.000	1.000	1.000	0.053	0.000

Table K.64 CPT corresponding to *outcome*

Estimated independence	*Times asked help*	
	'(-inf-0.5]'	'(0.5-inf)'
'(-inf-1.5]'	0.384	0.735
'(1.5-inf)'	0.616	0.265

Focus coarse value	*Interval of interaction*	
	'(773-inf)'	'(-inf-773]'
'(-inf-2.55]'	0.628	0.377
'(2.55-inf)'	0.372	0.623

Table K.65 CPTs of *estimated independence* and *focus coarse value*

Control		Irrelevant						
Type of outcome		PositiveAcceleration	Black-Out	Success	Distance-TooFar	Distance-TooShort	TimeOut	Re-Started
Publishing outcome	Unpublished	1.000	1.000	0.467	1.000	1.000	1.000	0.500
	Published	0.000	0.000	0.533	0.000	0.000	0.000	0.500

Table K.66 CPT corresponding to *publishing outcome* part 1

Control		Other						
Type of outcome		PositiveAcceleration	Black-Out	Success	Distance-TooFar	Distance-TooShort	TimeOut	Re-Started
Publishing outcome	Unpublished	1.000	1.000	0.731	1.000	1.000	1.000	1.000
	Published	0.000	0.000	0.269	0.000	0.000	0.000	0.000

Table K.67 CPT corresponding to *publishing outcome* part 2

Control		Self						
Type of outcome		PositiveAccel-eration	Black-Out	Success	Distance-TooFar	Distance-TooShort	TimeOut	Re-Started
Publish-ing out-come	Un-published	1.000	1.000	0.200	1.000	1.000	1.000	0.500
	Published	0.000	0.000	0.800	0.000	0.000	0.000	0.500

Table K.68 CPT corresponding to *publishing outcome* part 3

Interval of interaction		'(773-inf)'			'(-inf-773]'		
Average quality tutoring feedback		'(1.53-1.595]'	'(1.595-inf)'	'(-inf-1.53]'	'(1.53-1.595]'	'(1.595-inf)'	'(-inf-1.53]'
Times asked help	'(-inf-0.5]'	0.492	0.000	0.000	0.696	0.000	0.000
	'(0.5-inf)'	0.507	1.000	1.000	0.304	1.000	1.000

Table K.69 CPT corresponding to *times asked help*

Average quality of tutoring feedback	*Overall attempts*	
	'(-inf-2.5]'	'(2.5-inf)'
'(1.53-1.595]'	0.966	0.464
'(1.595-inf)'	0.017	0.193
'(-inf-1.53]'	0.017	0.343

Control t-1	
High	0.320
Low	0.228
Other	0.197
Irrelevant	0.108
Self	0.104
Medium	0.042

Table K.70 CPT of *average quality of tutoring feedback and control t-1*

Interval of interaction		'(773-inf)'		'(-inf-773]'	
Time to achieve learning goals		"(1981.664-inf)"	"(-inf-1981.664]"	"(1981.664-inf)"	"(-inf-1981.664]"
Type of outcome	PositiveAcceleration	0.078	0.000	0.594	0.034
	BlackOut	0.222	0.077	0.275	0.000
	Success	0.167	0.589	0.000	0.965
	DistanceTooFar	0.100	0.051	0.039	0.000
	DistanceTooShort	0.244	0.282	0.079	0.000
	TimeOut	0.178	0.000	0.029	0.000
	ReStarted	0.011	0.000	0.000	0.000

Table K.71 CPT corresponding to *type of outcome or end condition*

Estimated independence		'(-inf-1.5]'		'(1.5-inf)'	
Overall attempts		'(-inf-2.5]'	'(2.5-inf)'	'(-inf-2.5]'	'(2.5-inf)'
Attempts alone	'(-inf-2.5]'	1.000	0.462	1.000	0.000
	'(2.5-inf)'	0.000	0.538	0.000	1.000

Table K.72 CPT corresponding to *attempts alone*

Type of outcome		PositiveAcceleration		BlackOut		Success	
Interval of interaction		'(773-inf)'	'(-inf-773]'	'(773-inf)'	'(-inf-773]'	'(773-inf)'	'(-inf-773]'
Overall at-tempts	'(-inf-2.5]'	0.286	0.934	0.130	0.385	0.105	0.750
	'(2.5-inf)'	0.714	0.065	0.869	0.615	0.895	0.250

Table K.73 CPT corresponding to *overall attempts* part 1

Type of outcome		DistanceTooFar		DistanceTooShort		TimeOut		ReStarted	
Interval of interaction		'(773-inf)'	'(-inf-773]'	'(773-inf)'	'(-inf-773]'	'(773-inf)'	'(-inf-773]'	'(773-inf)'	'(-inf-773]'
Overall attempts	'(-inf-2.5]'	0.364	0.750	0.091	0.750	0.250	0.667	0.000	0.500
	'(2.5-inf)'	0.636	0.250	0.909	0.250	0.750	0.333	1.000	0.500

Table K.74 CPT corresponding to *overall attempts* part 2

Value t-1	*Control t-1*					
	High	Low	Other	Irrelevant	Self	Medium
Positive	0.723	0.000	0.059	0.286	0.296	0.818
Negative	0.277	0.508	0.941	0.714	0.704	0.181
None	0.000	0.492	0.000	0.000	0.000	0.000

Table K.75 CPT corresponding to *value t-1*

Control		Irrelevant		Other		Self	
Value		Negative	Positive	Negative	Positive	Negative	Positive
Emotion	Sadness	1.000	0.000	0.000	0.000	0.000	0.000
	Anger	0.000	0.000	1.000	0.000	0.000	0.000
	Joy	0.000	1.000	0.000	0.000	0.000	0.000
	Pride	0.000	0.000	0.000	0.000	0.000	1.000
	Shame	0.000	0.000	0.000	0.000	1.000	0.000
	Gartitude	0.000	0.000	0.000	1.000	0.000	0.000

Table K.76 CPT corresponding to *retrospective outcome emotions*

Average quality tutoring feedback		'(1.53-1.595]'					
Control t-1		High	Low	Other	Irrelevant	Self	Medium
Control	Irrelevant	0.317	0.342	0.100	0.591	0.263	0.454
	Other	0.400	0.368	0.733	0.136	0.263	0.454
	Self	0.283	1.289	0.167	0.273	0.474	0.091

Table K.77 CPT corresponding to *control* part 1

Average quality tutoring feedback		'(1.595-inf)'					
Control t-1		High	Low	Other	Irrelevant	Self	Medium
Control	Irrelevant	0.000	0.444	0.250	0.667	0.000	0.333
	Other	0.700	0.222	0.750	0.333	0.333	0.333
	Self	0.300	0.333	0.000	0.000	0.667	0.333

Table K.78 CPT corresponding to *control* part 2

Average quality tutoring feedback		'(-inf-1.53]'					
Control t-1		High	Low	Other	Irrelevant	Self	Medium
Control	Irrelevant	0.077	0.167	0.000	0.000	0.200	0.333
	Other	0.538	0.667	0.882	0.333	0.000	0.333
	Self	0.385	0.167	0.118	0.667	0.800	0.333

Table K.79 CPT corresponding to *control* part 3

Control t-1		High											
Type of outcome		PositiveAcceleration						BlackOut					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.917	1.000	0.500	0.500	0.500	0.500	0.700	1.000	0.500	0.500	0.500	0.500
	Positive	0.083	0.000	0.500	0.500	0.500	0.500	0.300	0.000	0.500	0.500	0.500	0.500

Table K.80 CPT corresponding to *value* part 1

Control t-1		High											
Type of outcome		Success						DistanceTooFar					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.600	0.667	0.500	0.154	0.000	0.500	1.000	1.000	0.500	0.500	0.500	0.500
	Positive	0.400	0.333	0.500	0.846	1.000	0.500	0.000	0.000	0.500	0.500	0.500	0.500

Table K.81 CPT corresponding to *value* part 2

Control t-1		High											
Type of outcome		DistanceTooShort						TimeOut					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.778	0.833	0.500	0.500	0.500	0.500	0.667	1.000	0.500	0.500	0.500	0.500
	Positive	0.222	0.167	0.500	0.500	0.500	0.500	0.333	0.000	0.500	0.500	0.500	0.500

Table K.82 CPT corresponding to *value* part 3

Control t-1		High						Low					
Type of outcome		ReStarted						PositiveAcceleration					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	1.000	0.500	0.500	0.500	0.500	0.500	0.846	0.889	0.500	0.500	0.500
	Positive	0.500	0.000	0.500	0.500	0.500	0.500	0.500	0.154	0.111	0.500	0.500	0.500

Table K.83 CPT corresponding to *value* part 4

Control t-1		Low											
Type of outcome		BlackOut						Success					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	0.667	0.800	0.500	0.500	0.500	0.500	0.500	0.400	0.500	0.167	0.000
	Positive	0.500	0.333	0.200	0.500	0.500	0.500	0.500	0.500	0.600	0.500	0.833	1.000

Table K.84 CPT corresponding to *value* part 5

Control t-1		Low											
Type of outcome		DistanceTooFar						DistanceTooShort					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	1.000	1.000	0.500	0.500	0.500	0.500	1.000	0.500	0.500	0.500	0.500
	Positive	0.500	0.000	0.000	0.500	0.500	0.500	0.500	0.000	0.500	0.500	0.500	0.500

Table K.85 CPT corresponding to *value* part 6

Control t-1		Low											
Type of outcome		TimeOut						ReStarted					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	1.000	1.000	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	Positive	0.500	0.000	0.000	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500

Table K.86 CPT corresponding to *value* part 7

Control t-1		Other											
Type of outcome		PositiveAcceleration						BlackOut					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	0.875	0.500	0.500	0.500	0.500	0.500	0.929	0.500	0.500	0.500	0.500
	Positive	0.500	0.125	0.500	0.500	0.500	0.500	0.500	0.071	0.500	0.500	0.500	0.500

Table K.87 CPT corresponding to *value* part 8

Control t-1		Other											
Type of outcome		Success						DistanceTooFar					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	1.000	0.500	0.000	0.200	0.500	0.500	1.000	0.500	0.500	0.500	0.500
	Positive	0.500	0.000	0.500	1.000	0.800	0.500	0.500	0.000	0.500	0.500	0.500	0.500

Table K.88 CPT corresponding to *value* part 9

Control t-1		Other						Irrelevant					
Type of outcome		ReStarted						PositiveAcceleration					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	0.500	0.500	0.500	0.500	0.500	1.000	1.000	0.500	0.500	0.500	0.500
	Positive	0.500	0.500	0.500	0.500	0.500	0.500	0.000	0.000	0.500	0.500	0.500	0.500

Table K.89 CPT corresponding to *value* part 10

Control t-1		Irrelevant											
Type of outcome		BlackOut						Success					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.000	1.000	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.000	0.000	0.500
	Positive	1.000	0.000	0.500	0.500	0.500	0.500	0.500	0.500	0.500	1.000	1.000	0.500

Table K.90 CPT corresponding to *value* part 11

Control t-1		Irrelevant											
Type of outcome		DistanceTooFar						DistanceTooShort					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	1.000	0.500	0.500	0.500	0.500	0.500	0.333	0.500	0.500	0.500	0.500
	Positive	0.500	0.000	0.500	0.500	0.500	0.500	0.500	0.667	0.500	0.500	0.500	0.500

Table K.91 CPT corresponding to *value* part 12

Control t-1		Irrelevant											
Type of outcome		TimeOut						ReStarted					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	1.000	1.000	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	Positive	0.000	0.000	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500

Table K.92 CPT corresponding to *value* part 13

Control t-1		Self											
Type of outcome		PositiveAcceleration						BlackOut					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	1.000	1.000	0.500	0.500	0.500	0.500	0.000	1.000	0.500	0.500	0.500	0.500
	Positive	0.000	0.000	0.500	0.500	0.500	0.500	1.000	0.000	0.500	0.500	0.500	0.500

Table K.93 CPT corresponding to *value* part 14

Control t-1		Self											
Type of outcome		Success						DistanceTooFar					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	0.500	0.500	0.500	0.667	0.500	1.000	1.000	0.500	0.500	0.500	0.500
	Positive	0.500	0.500	0.500	0.500	0.333	0.500	0.000	0.000	0.500	0.500	0.500	0.500

Table K.94 CPT corresponding to *value* part 15

Control t-1		Self											
Type of outcome		DistanceTooShort						TimeOut					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	0.500	0.500	0.500	0.667	0.500	1.000	1.000	0.500	0.500	0.500	0.500
	Positive	0.500	0.500	0.500	0.500	0.333	0.500	0.000	0.000	0.500	0.500	0.500	0.500

Table K.95 CPT corresponding to *value* part 16

Control t-1		Self						Medium					
Type of outcome		ReStarted						PositiveAcceleration					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	0.500	0.500	0.500	0.500	0.500	0.857	0.500	0.500	0.500	0.500	0.500
	Positive	0.500	0.500	0.500	0.500	0.500	0.500	0.143	0.500	0.500	0.500	0.500	0.500

Table K.96 CPT corresponding to *value* part 17

Control t-1		Medium											
Type of outcome		BlackOut						Success					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	0.500	0.500	0.500	0.500	0.500	1.000	0.500	0.500	0.500	0.500	0.500
	Positive	0.500	0.500	0.500	0.500	0.500	0.500	0.000	0.500	0.500	0.500	0.500	0.500

Table K.97 CPT corresponding to *value* part 18

Control t-1		Medium											
Type of outcome		DistanceTooFar						DistanceTooShort					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	0.500	0.500	0.500	0.500	0.500	1.000	0.500	0.500	0.500	0.500	0.500
	Positive	0.500	0.500	0.500	0.500	0.500	0.500	0.000	0.500	0.500	0.500	0.500	0.500

Table K.98 CPT corresponding to *value* part 19

Control t-1		Medium											
Type of outcome		TimeOut						ReStarted					
Publishing outcome		Unpublished			Published			Unpublished			Published'		
Value t-1		Positive	Negative	None	Positive	Negative	None	Positive	Negative	None	Positive	Negative	None
Value	Negative	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
	Positive	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500

Table K.99 CPT corresponding to *value* part 20

K.9 Pearson correlations of activity outcome emotions variables including GSR

The Pearson correlations for the variables involved in the creation of the model corresponding to the activity outcome emotions are presented for 46 cases.

	Times asked help	Attempts alone	Estimated independence	Overall attempts	Interval of interaction	Focus coarse value	GSR signal (2 cat)	GSR signal (3 cat)	Emotion	Control	Value	Control t-1	Value t-1
Times asked help													
Pearson correlations	1	0.269	0.147	0.269	-0.258	0.387**	-0.387**	-0.320*	0.152	0.199	-0.405**	-0.017	-0.259
Sig. (2-tailed)		0.070	0.329	0.070	0.083	0.008	0.008	0.030	0.314	0.184	0.005	0.910	0.082
Attempts alone													
Pearson correlations	0.269	1	0.953**	1.000**	0.325*	0.325*	-0.325*	-0.402**	0.079	-0.170	-0.258	-0.323*	-0.291*
Sig. (2-tailed)	0.070		2.39E-24	2.39E-24	0.028	0.028	0.028	0.006	0.600	0.259	0.083	0.028	0.049
Estimated independence													
Pearson correlations	0.147	0.953**	1	0.953**	0.274	0.274	-0.274	-0.339*	0.017	-0.222	-0.206	-0.270	-0.249
Sig. (2-tailed)	0.329	2.39E-24		2.39E-24	0.066	0.066	0.066	0.021	0.911	0.139	0.169	0.070	0.097
Overall attempts													
Pearson correlations	0.269	1.000**	0.953**	1	0.325*	0.325*	-0.325*	-0.402**	0.079	-0.170	-0.258	-0.323*	-0.291*
Sig. (2-tailed)	0.070	2.39E-24	2.39E-24		0.028	0.028	0.028	0.006	0.600	0.259	0.083	0.028	0.049

Table K.100 Pearson correlations of activity emotion variables including GSR part 1

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'.

	Times asked help	Attempts alone	Estimated independence	Overall attempts	Interval of interaction	Focus coarse value	GSR signal (2 cat)	GSR signal (3 cat)	Emotion	Control	Value	Control t-1	Value t-1
Interval of interaction													
Pearson correlations	-0.258	0.325*	0.274	0.325*	1	0.043	-0.130	-0.054	-0.062	-0.044	-0.365*	-0.033	-0.416**
Sig. (2-tailed)	0.083	0.028	0.066	0.028		0.774	0.388	0.722	0.683	0.771	0.013	0.827	0.004
Focus coarse value													
Pearson correlations	0.387**	0.325*	0.274	0.325*	0.043	1	-0.478**	-0.215	-0.062	-0.044	-0.073	0.066	-0.269
Sig. (2-tailed)	0.008	0.028	0.066	0.028	0.774		0.001	0.151	0.683	0.771	0.630	0.661	0.071
GSR signal (2 cat)													
Pearson correlations	-0.387**	-0.325*	-0.274	-0.325*	-0.130	-0.478**	1	0.808**	-0.021	-0.132	0.267	-0.100	0.318*
Sig. (2-tailed)	0.008	0.028	0.066	0.028	0.388	0.001		1.19E-11	0.892	0.380	0.072	0.510	0.031
GSR signal (3 cat)													
Pearson correlations	-0.320*	-0.402**	-0.339*	-0.402**	-0.054	-0.215	0.808**	1	-0.153	0.000	0.090	0.082	0.182
Sig. (2-tailed)	0.030	0.006	0.021	0.006	0.722	0.151	1.19E-11		0.310	1.000	0.551	0.587	0.227
Emotion													
Pearson correlations	0.152	0.079	0.017	0.079	-0.062	-0.062	-0.021	-0.153	1	0.370*	0.202	-0.056	0.173
Sig. (2-tailed)	0.314	0.600	0.911	0.600	0.683	0.683	0.892	0.310		0.011	0.178	0.711	0.251

Table K.101 Pearson correlations of activity emotion variables including GSR part 2

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'.

	Times asked help	Attempts alone	Estimated independence	Overall attempts	Interval of interaction	Focus coarse value	GSR signal (2 cat)	GSR signal (3 cat)	Emotion	Control	Value	Control t-1	Value t-1
Control													
Pearson correlations	0.199	-0.170	-0.222	-0.170	-0.044	-0.044	-0.132	0.000	0.370*	1	-0.432**	-0.015	-0.172
Sig. (2-tailed)	0.184	0.259	0.139	0.259	0.771	0.771	0.380	1.000	0.011		0.003	0.923	0.254
Value													
Pearson correlations	-0.405**	-0.258	-0.206	-0.258	-0.365*	-0.073	0.267	0.090	0.202	-0.432**	1	0.040	0.582**
Sig. (2-tailed)	0.005	0.083	0.169	0.083	0.013	0.630	0.072	0.551	0.178	0.003		0.790	2.24E-5
Control t-1													
Pearson correlations	-0.017	-0.323*	-0.270	-0.323*	-0.033	0.066	-0.100	0.082	-0.056	-0.015	0.040	1	-0.312*
Sig. (2-tailed)	0.910	0.028	0.070	0.028	0.827	0.661	0.510	0.587	0.711	0.923	0.790		0.035
Value t-1													
Pearson correlations	-0.259	-0.291*	-0.249	-0.291*	-0.416**	-0.269	0.318*	0.182	0.173	-0.172	0.582**	-0.312*	1
Sig. (2-tailed)	0.082	0.049	0.097	0.049	0.004	0.071	0.031	0.227	0.251	0.254	2.24E-5	0.035	

Table K.102 Pearson correlations of activity emotion variables including GSR part 3

Correlations that are significant at 0.01 level (2-tailed) are signalled with '**', whist correlations that are significant at 0.05 level (2-tailed) are signalled with '*'.

K.10 CPTs corresponding to activity outcome emotions network including GSR

Interval of interaction	*Attempts alone*	
	"(-inf-0.5]"	"(0.5-inf)"
'(-inf-380.5]'	0.733	0.387
'(380.5-inf)'	0.267	0.613

Overall attempts	*Attempts alone*	
	"(-inf-0.5]"	"(0.5-inf)"
'(-inf-0.5]'	1.000	0.000
'(0.5-inf)'	0.000	1.000

Control	*Value t-1*		
	Positive	None	Negative
Low	0.214	0.714	0.44
High	0.786	0.286	0.56

Table K.103 CPTs of *interval of interaction, overall attempts* and *control*

Estimated independence	
'(-inf-0.5]''	0.348
'(0.5-inf)'	0.652

Value t-1	
Positive	0.304
None	0.152
Negative	0.543

Table K.104 CPTs corresponding to *estimated independence* and *value t-1*

Estimated independence		'(-inf-0.5]'			'(0.5-inf)'		
Value t-1		Positive	None	Negative	Positive	None	Negative
Control t-1	High	0.500	0.000	0.200	1.000	0.000	0.350
	Low	0.000	1.000	0.200	0.000	1.000	0.400
	Medium	0.500	0.000	0.600	0.000	0.000	0.000
	Irrelevant	0.000	0.000	0.000	0.000	0.000	0.100
	Self	0.000	0.000	0.000	0.000	0.000	0.150

Table K.105 CPT corresponding to *control t-1*

Estimated independence		'(-inf-0.5]'			'(0.5-inf)'		
Value		None	Negative	Positive	None	Negative	Positive
Attempts alone	'(-inf-0.5]'	1.000	0.667	1.000	0.000	0.000	0.000
	'(0.5-inf)'	0.000	0.333	0.000	1.000	1.000	1.000

Table K.106 CPT corresponding to *attempts alone*

Control t -1		High			Low			Medium		
Value t – 1		Positive	None	Negative	Positive	None	Negative	Positive	None	Negative
Value	None	0.182	0.333	0.125	0.333	0.571	0.000	0.000	0.333	0.667
	Negative	0.091	0.333	0.875	0.333	0.286	1.000	0.000	0.333	0.000
	Positive	0.727	0.333	0.000	0.333	0.143	0.000	1.000	0.333	0.333

Table K.107 CPT *corresponding to value* part 1

Control t -1		Irrelevant			Self		
Value t – 1		Positive	None	Negative	Positive	None	Negative
Value	None	0.333	0.333	0.000	0.333	0.333	0.000
	Negative	0.333	0.333	0.000	0.333	0.333	0.333
	Positive	0.333	0.333	1.000	0.333	0.333	0.667

Table K.108 CPT *corresponding to value* part 2

Control t -1		Low			High		
Value t – 1		None	Negative	Positive	None	Negative	Positive
Emotion	Boredom	1.000	0.000	0.250	0.250	0.000	0.000
	Frustration	0.000	1.000	0.250	0.250	0.000	0.000
	Anger	0.000	0.000	0.250	0.250	1.000	0.000
	Enjoyment	0.000	0.000	0.250	0.250	0.000	1.000

Table K.109 CPT *corresponding to activity emotions*

K.11 CPTs corresponding to activity emotions network excluding GSR

Overall attempts	*Attempts alone*	
	"(-inf-0.5]"	"(0.5-inf)"
'(-inf-0.5]'	1.000	0.000
'(0.5-inf)'	0.000	1.000

Attempts alone	
'(-inf-0.5]"	0.326
'(0.5-inf)'	0.674

Estimated independence	
'(-inf-0.5]"	0.348
'(0.5-inf)'	0.652

Table K.110 CPTs of *overall attempts, attempts alone* and *estimated independence*

Control	*Value t-1*		
	Positive	None	*Negative*
Low	0.214	0.714	0.440
High	0.786	0.286	0.560

Value t-1	
Positive	*0.304*
None	0.152
Negative	0.543

Table K.111 CPTs of *control* and *value t-1*

Estimated independence		'(-inf-0.5]'			'(0.5-inf)'		
Value t-1		Positive	None	Negative	Positive	None	Negative
Control t-1	High	0.500	0.000	0.200	1.000	0.000	0.350
	Low	0.000	1.000	0.200	0.000	1.000	0.400
	Medium	0.500	0.000	0.600	0.000	0.000	0.000
	Irrelevant	0.000	0.000	0.000	0.000	0.000	0.100
	Self	0.000	0.000	0.000	0.000	0.000	0.150

Table K.112 CPT corresponding to *control t-1*

Attempts alone		'(-inf-0.5]'			'(0.5-inf)'		
Value t-1		Positive	None	Negative	Positive	None	Negative
Value	None	0.167	0.800	0.500	0.125	0.000	0.048
	Negative	0.167	0.000	0.250	0.000	1.000	0.762
	Positive	0.666	0.200	0.250	0.875	0.000	0.190

Table K.113 CPT corresponding to *value*

Control t -1		Low			High		
Value t -1		None	Negative	Positive	None	Negative	Positive
Emotion	Boredom	1.000	0.000	0.250	0.250	0.000	0.000
	Frustration	0.000	1.000	0.250	0.250	0.000	0.000
	Anger	0.000	0.000	0.250	0.250	1.000	0.000
	Enjoyment	0.000	0.000	0.250	0.250	0.000	1.000

Table K.114 CPT corresponding to *activity emotions*

K.12 Evaluation of PlayPhysics GBL environment

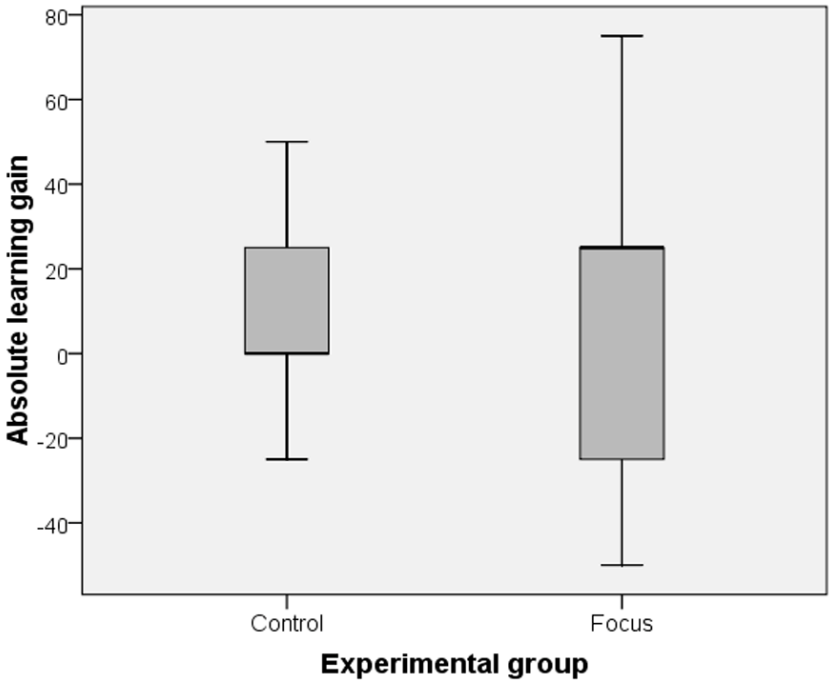


Figure K.1 Boxplots of absolute learning gain after removing outliers

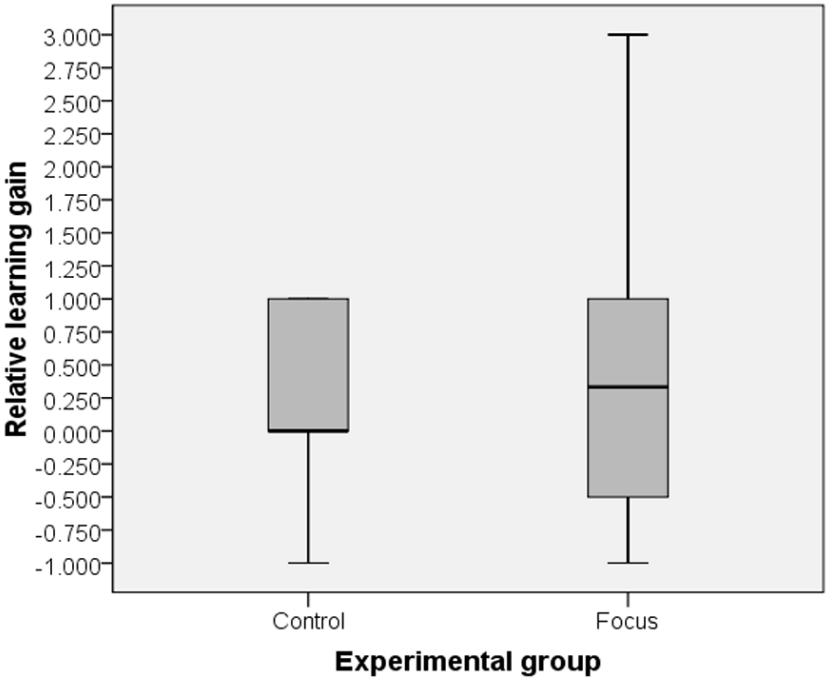


Figure K.2 Boxplots of relative learning gain after removing extreme cases and outliers

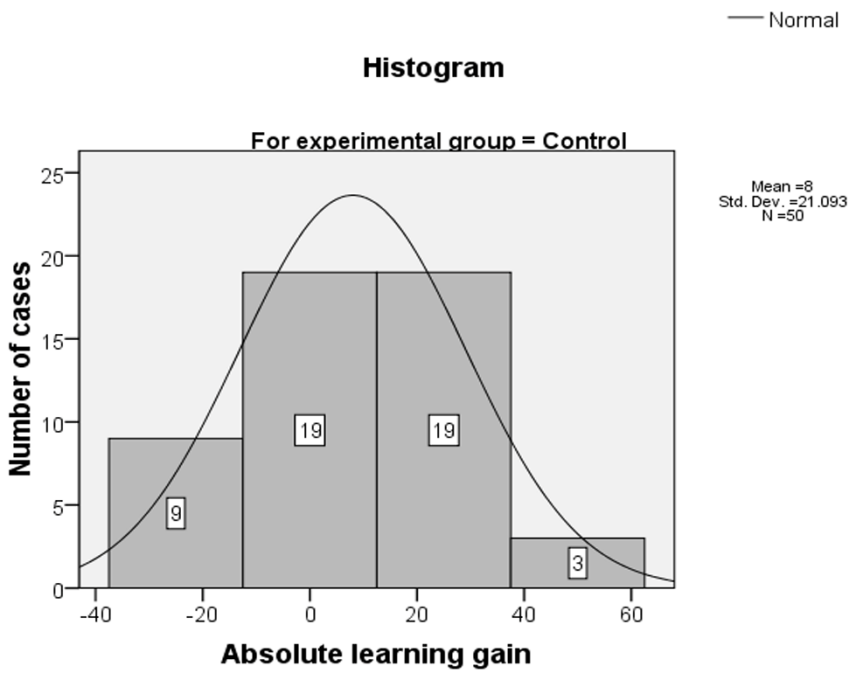


Figure K.3 Histogram of the control group absolute learning gain

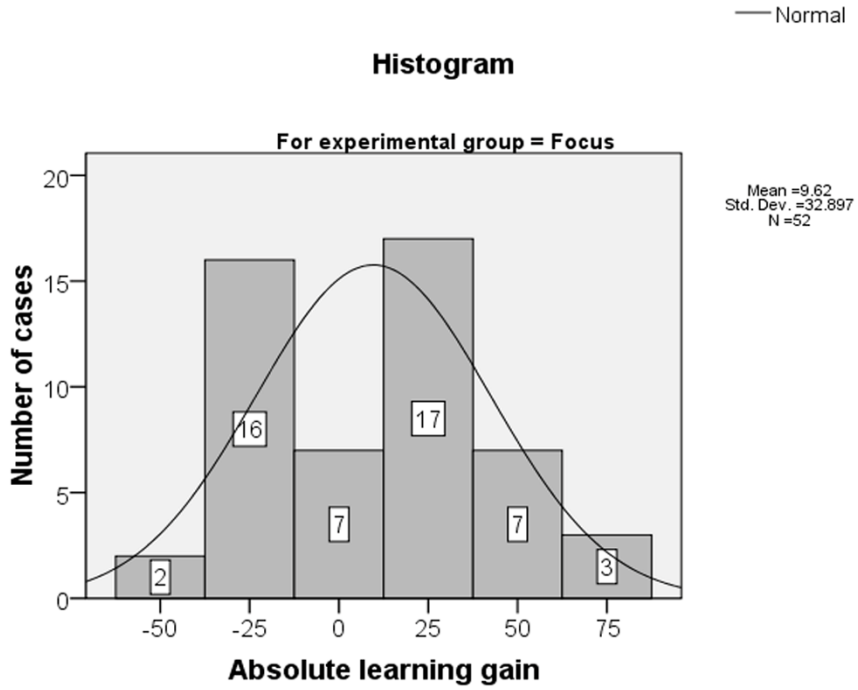


Figure K.4 Histogram of the focus group absolute learning gain

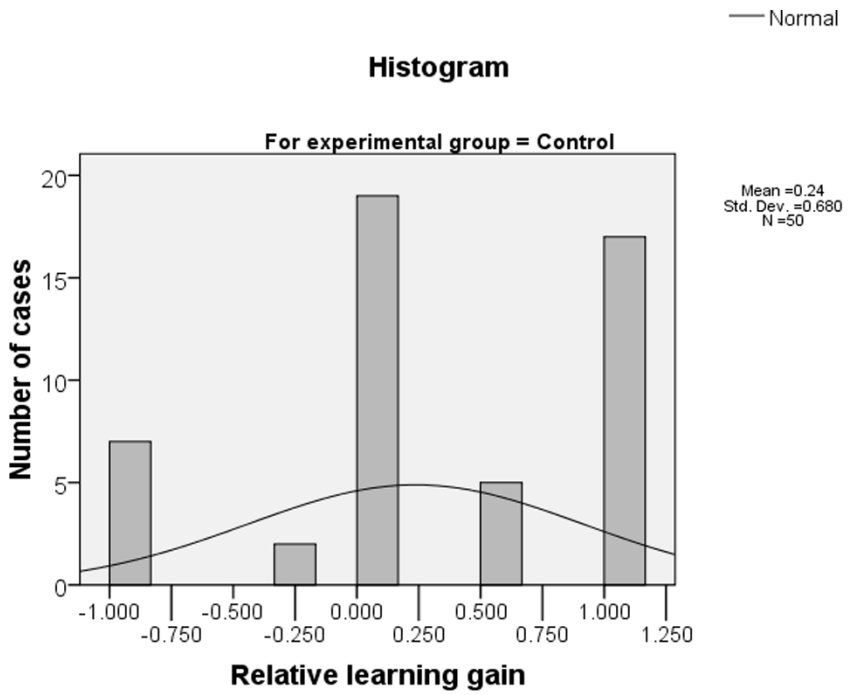


Figure K.5 Histogram of control group relative learning gain

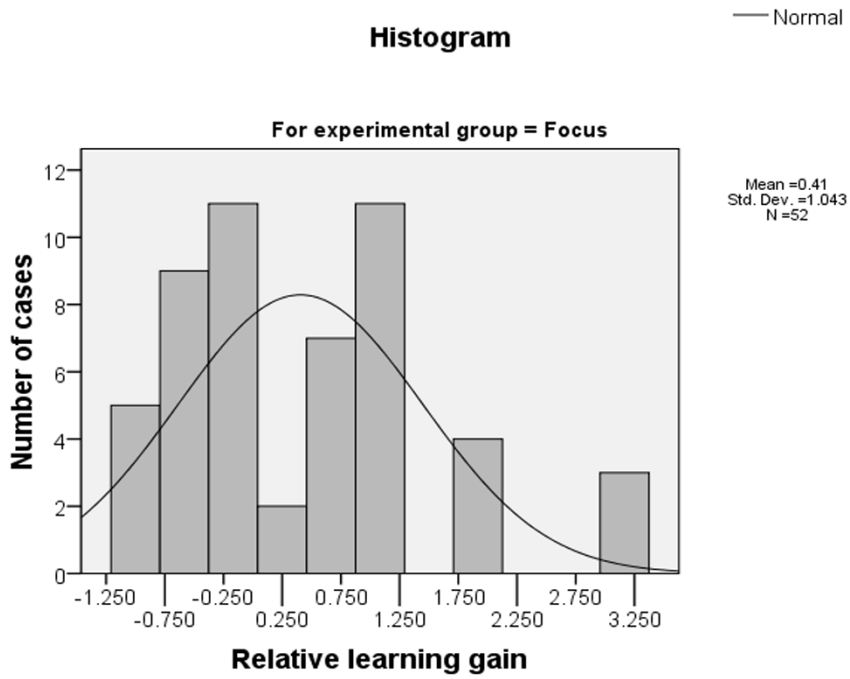


Figure K.6 Histogram corresponding to the relative learning gain of the focus group

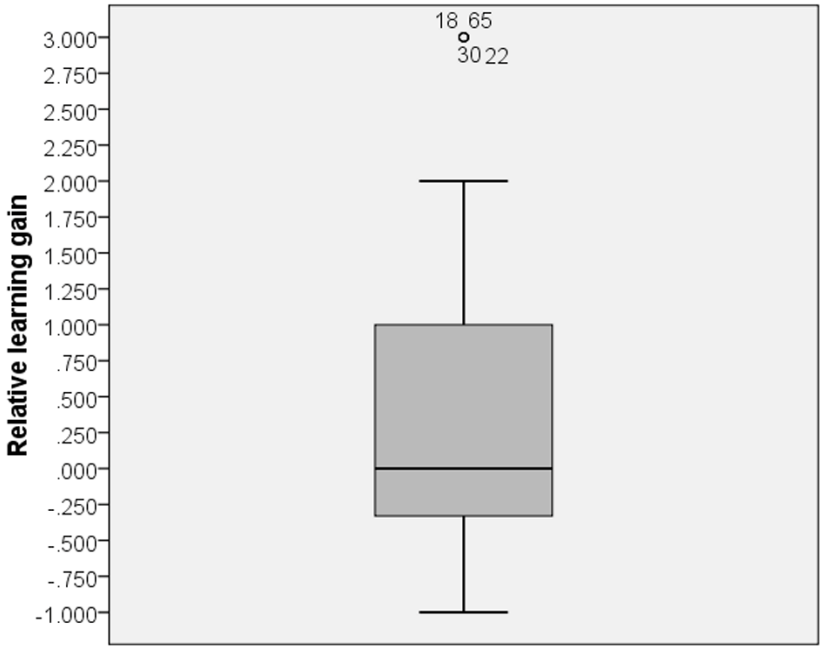


Figure K.7 Boxplot of the relative learning gain corresponding to 104 students

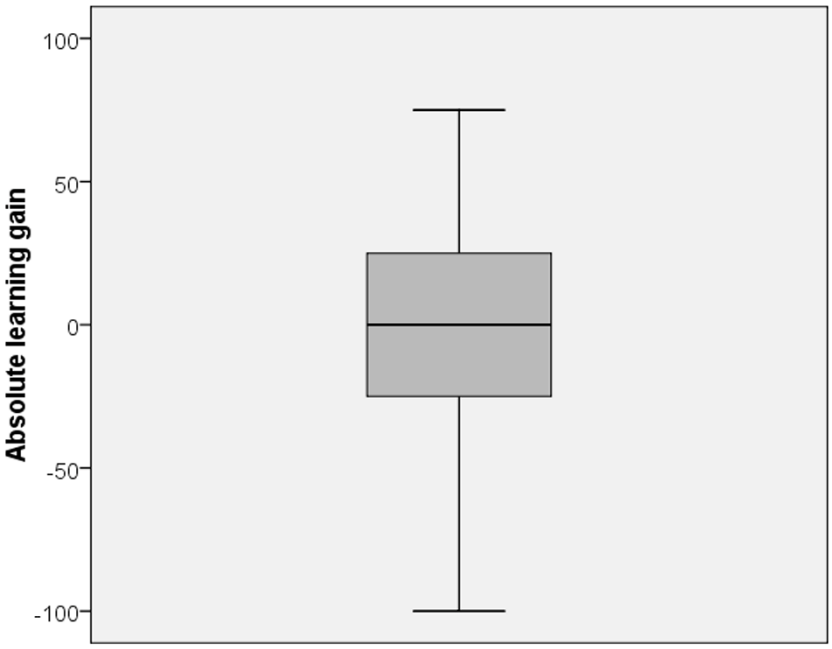


Figure K.8 Boxplot of the absolute learning gain corresponding to 104 students

References

Adobe (2012) 'Adobe Creative Suite Family', *Website Design Software* [online], available: http://www.adobe.com/products/creativesuite.html?promoid=JOLIS [accessed 27 March 2012].

Affectiva (2012) 'Q Sensor: Respectful Emotion Measurement and Communication', *Q Sensor 2.0* [online], available: http://www.affectiva.com/q-sensor/ [accessed 27 July 2012].

Ahmad, N. and Bull, S. (2008) 'Do Students Trust Their Open Learner Models?', in Nejd, W., Kay, J., Pu, P. and Herder, E., eds., *Proc. of the 5th International Conference on Adaptive Hypermedia and Adapted Web-Based Systems (AH 2008)*, Hannover, Germany, 29 July - 1 August, Berlin Heidelberg: Springer-Verlag, 255-258.

Ajzen, I. (2005) *Attitudes, Personality and Behaviour* 2ed., Maidenhead, Berkshire, England: Open University Press.

Akilli, G. K. and Cagiltay, K. (2006) 'An Instructional Design/Development Model for the Creation of Game-Like Learning Environments: The FIDGE Model' in Pivec, M., ed. *Affective and Emotional Aspects of Human-Computer Interaction*, Netherlands, Amsterdam: IOS Press, 93-111.

Alea, N., Dielh, M. and Bluck, S. (2004) *Personality and Emotion in Late Life, Encyclopedia of Applied Psychology,* San Diego, CA: Elsevier.

Alexander, S. and Sarrafzadeh, A. (2008) 'Foundation of An Affective Tutoring System: Learning How Human Tutors Adapt to Student Emotion', *International Journal of Intelligent Systems Technologies and Applications,* 4(3/4), 355-367.

Alexander, S., Sarrafzadeh, A. and Hill, S. (2006) 'Easy with Eve: A Functional Affective Tutoring System', in Rebolledo-Mendez, G. and Martinez-Miron, E., eds., *Proc. of 8th International Conference on Intelligent Tutoring Systems (ITS 2006): Workshop on Motivational and Affective Issues in ITS*, Jhongli, Taiwan, 26-30 June 2006, 38-45.

Anderson, J. R. and Skwarecki, E. (1986) 'The Automated Tutoring of Introductory Computing Programming ', *Communications of the ACM*, 842-849.

Andersson, G., Höök, K., Mourão, D., Paiva, A. and Costa, M. (2002) 'Using a Wizard of Oz Study to Inform the Design of SenToy', in *Proc. of the 4th International Conference on Designing Interactive Systems (DIS-02): processes, practices, methods and techniques*, London, 25th - 28th June, New York, NY, USA: ACM Press, 349-355.

Andrade, A. F., Vicari, R. M., Jung, J. L., Bordini, R. and Jaques, P. A. (2001) 'A Computational Model of Distance Learning based on Vygotsky's Socio-cultural Approach', in Vassileva, J., ed. *Proc of the International Conference on Artificial Intelligence in Education*

(AIED 2001): the Workshop Multi-Agent Architectures for Distributed Learning Environments, San Antonio, TX, USA, 19-23 May 2001, Saskatoon, Saskatchewan, Canada: IAIED Society, 33-40.

Apache Software Foundation (2011) 'LOG4J', *Logging Services* [online], available: http://logging.apache.org/log4j/1.2/ [accessed 27 March 2011].

Aqel, M. M., Mahanti, P. K. and Banerjee, S. (2009) 'Analyzing Artificial Emotion in Game Characters Using Soft Computing', *Engineering and Technology,* 53, 926-929.

Arroyo, I., Cooper, D. G., Burleson, W., Woolf, B. P., Muldner, C. and Christopherson, R. (2009) 'Emotion Sensors Go to School', in Dimitrova, V., Mizoguchi, R., Du Boulay, B. and Graesser, A., eds., *Proc. of the 14th International Conference on Artificial Intelligence in Education (AIED 2009): Building Learning Systems that Care: From Knowledge Representation to Affective Modelling*, Brighton, UK, 6-10 July 2009, Amsterdam, The Netherlands: IOS Press, 17-24.

Arroyo, I. and Woolf, B. P. (2005) 'Inferring learning and attitudes from a Bayesian network of log file data' in Looi, C. K., Mc Calla, G., Bredeweg, B. and Breuker, J., eds., *Proc. of the 12th International Conference on Artificial Intelligence in Education, Frontiers in Artificial Intelligence and Applications*, Amsterdam: IOS Press, 33-40.

Autodesk (2012) 'Autodesk 3ds Max Products', *3D Modelling, Animation and Rendering Software* [online], available: http://usa.autodesk.com/3ds-max/ [accessed 27 March 2012].

Bakkes, S., Tan , C. T. and Pisan, Y. (2012) 'Personalised Gaming: A motivation and Overview of Literature', in Cermak-Sassenrath, D., Walker, C. and Tan, C. T., eds., *Proc. of the 8th Australasian Conference on Interactive Entertainment (IE 2012) : Playing the System*, Auckland, NZ, New Zealand, 21 - 22 July 2012, New York, NY: ACM, 4:1-4:10.

Bandura, A. (1995) *Self-Efficacy in Changing Societies,* Cambridge: Cambridge University Press.

Bandura, A. (1997) *Self-efficacy: The Excercise of Control,* New York, NY: Freeman.

Bandura, A. (2006) 'Guide for Constructing Self-efficacy Scales' in Pajares, F. and Urdan, T., eds., *Self-Efficacy Beliefs of Adolescents*, Greenwich, Connecticut: Information Age Publishing, 307-337.

Bashar, A., Parr, G., McClean, S., Scotney, B. and Nauck, D. (2010) *Learning-Based Call Admission Control Framework for QoS Management in Heterogeneous Networks,* translated by Zavoral, F., Yaghob, J., Pichappan, P. and El-Qawasmeh, E., Prague, Czech Republic: Springer-Verlag, 99-111.

Bateman, C. and Boon, R. (2006) *21st Century Game Design,* Hingham, Massachussets: Charles River Media, Inc.

Bayne, R. (1997) *The Myers-Briggs type indicator : a critical review and practical guide,* Cheltenham, Gluocestershire: Stanley Thornes.

Becker-Asano, C. and Wachsmuth, I. (2010) 'Affective Computing with Primary and Secondary Emotions in a Virtual Human', in Kopp, S., Aylett, R., Gratch, J., Oliver, P. L. and Pelachaud, C., eds., *Proc. of the 9th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS10)*, Toronto, Canada, 10-14 May, Berlin: Springer, 32-49.

Berlyne, D. E. (1960) *Conflict, arousal and curiosity.,* New York: McGraw-Hill.

Bertram, D. (2012) 'Likert Scales', *Univerzitet u Beogradu - Matematčki fakultet - CPSC 681 - Topic Report* [online], available: http://poincare.matf.bg.ac.rs/~kristina//topic-dane-likert.pdf [accessed 29 July 2012].

Block, B. (2008) *The Visual Story: Creating the Visual Structure of Film, TV and Digital Media,* 2 ed.*,* Burlington, MA: Elsevier.

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.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H. and Krathwohl, D. R. N. Y., Longmans, Green, 1956. (1984) *Taxonomy of Educational Objectives: The Classification of Educational Goals Handbook I: Cognitive Domain Bloom's Taxonomy: A Forty-Year Retrospective,* New York: Longman.

Bouckaert, R., R. (2008) *Bayesian Network Classifiers in WEKA for version 3-5-7,* The University of Waikato [online], available: http://www.cs.waikato.ac.nz/~remco/weka.bn.pdf [accessed 20 July 2012].

Bouckaert, R., R., Frank, E., Hall, M., Kirkby, R., Reutemann, P., Seewald, A. and Sucuse, D. (2012) *WEKA Manual for version 3-7-6,* The University of Waikato [online], available: http://switch.dl.sourceforge.net/project/weka/documentation/3.7.x/WekaManual-3-7-6.pdf [accessed 20 July 2012].

Boud, D. and Feletti, G. (1997) *The Challenge of Problem-Based Learning,* 2 ed.*,* London: Kogan Page.

Brave, S. and Nass, C. (2008) 'Emotion in Human-Computer Interaction' in Sears, A. and Jacko, J. A., eds., *The Human Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, 2 ed., New York: Lawrence Earlbaum Associates, Taylor & Francis Group, 77-92.

Breazeal, C. and Brooks, R. (2005) 'Robot Emotion: A Functional Perspective' in Fellous, J.-M. and Arbib, M. A., eds., *Who Needs Emotions?: The Brain Meets the Robot*, New York, NY: Oxford University Press, 271-310.

Brown, P. and Levinson, S. (1987) 'Politness. Some Universals in Language Usage' in Gumperz, J. J., ed. *Studies in Interactional Sociolinguistics*, Cambridge, UK: Cambridge University Press.

Bruner, J. S. (1960) *The Process of Education,* Cambridge, Mass: Harvard University Press.

Brusilovsky, P., Kobsa, A. and Nejdl, W., eds. (2007) *The Adaptive Web: Methods and Strategies of Web Personalization, Lecturer Notes in Computer Science (LNCS) 4321,* Heidelberg, Berlin: Springer-Verlag.

Bunt, A. and Conati, C. (2003) 'Probabilistic Student Modeling to Improve Exploratory Behavior', *User Modeling and User-Adapted Interaction,* 13(3), 269-309. .

Burleson, W. (2006) *Affective Learning Companions: Strategies for Empathic Agents with Real Time Multimodal Affective Sensing to Foster Meta-Cognitive Approaches to Learning, Motivation and Perseverance*, unpublished thesis Massachusets Institute of Technology (MIT)

Burleson, W. and Picard, R. W. (2007) 'Evidence for Gender Specific Approaches to the Development of Emotionally Intelligent Learning Companions', *IEEE Intelligent Systems,* 22(4), 62-69.

Carpenter, A. and Windsor, H. (2006) 'A Head of the Game? - Games in Education', *Serious Games Source* [online], available: http://seriousgamessource.com/features/feature_061306_ahead_of_the_game.php. [accessed 24 January 2010].

Chalfoun, P., Chaffar, S. and Frasson, C. (2006) 'Predicting the Emotional Reaction of the Learner with a Machine Learning Technique', in Rebolledo-Mendez, G. and Martinez-Miron, E., eds., *Proc. of 8th International Conference on Intelligent Tutoring Systems (ITS 2006): Workshop on Motivational and Affective Issues in ITS*, Jhongli, Taiwan, 26-30 June 2006, 13-20.

Chumbley, J. and Griffiths, M. (2006) 'Affect and the Computer Game Player: The Effect of Gender, Personality, and the Game Reinforcement Structure on Affective Responses to Computer Game-Play', *CyberPsychology & Behaviour,* 9(3), 308-316.

Clancey, G. L. and Buchanan, B. (1982) *Exploration of Teaching and Problem Solving Strategies '1979 - 1982'*, Stanford, California, USA: Stanford University, unpublished.

Cohen, J. (1960) 'A Coefficient of Agreement for Nominal Scales', *Educational and Psychological Measurement,* 10, 37-46.

Cohen, J. (1988) *Statistical Power Analysis for the Behavioural Science,* 2 ed., Hillsdale, N.J.: Lawrence Erlbaum Associates.

Collins, K. (2008) *Game Sound: An Introduction to the History, Theory and Practice of Video Game Music and Sound Design,* Cambridge, MA: MIT Press.

Conati, C. (2002) 'Probabilistic Assessment of User's Emotions in Educational Games', *Applied Artificial Intelligence,* 16, 555-575.

Conati, C. (2004) 'How to Evaluate Models of User Affect?' in André, E., Dybkjær, L., Minker, W. and Heisterkamp, P., eds., *Affective Dialogue Systems*, Heidelberg, Berlin: Springer-Verlag, 288-300.

Conati, C. and Maclaren, H. (2009) 'Empirically Building and Evaluating a Probabilistic Model of User Affect.', *User Modeling and User-Adapted Interaction,* 19(3), 267-303.

Connolly, T. and Stansfield, M. (2006) 'Using Game-based Learning Technologies in Overcoming Difficulties in Teaching Information Systems', *Journal of Information Technology Education,* 5, 459-476.

Costa, M. and McCrae, R. R. (1992) 'Four Ways Five Factors are Basic', *Personality and Individual Differences,* 13, 653-665.

CPR Medical (2012) 'Pre-TENS: Conductive Skin Preparation', *Pre-TENS* [online], available: http://www.cprmedical.com/elec/Unipatch_Catalog_page10.htm [accessed 18 July 2012].

D'Mello, S., Craig, S., Witherspoon, A., McDaniel, B. T. and Graesser, A. (2008a) 'Automatic Detection of Learner's Affect from Conversational Cues', *User Modeling and User-Adapted Interaction,* 18(1-2), 45-80.

D'Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H., Person, N., Kort, B., el Kaliouby, R., Picard., R. W. and Graesser, A. (2008b) 'AutoTutor Detects and Responds to Learners Affective and Cognitive States ', in *International Conference of Intelligent Tutoring Systems, Workshop on Emotional and Cognitive Issues*, Montreal, Canada, 23 - 27 June,

D'Mello, S. K., Craig, S. D., Witherspoon, A., McDaniel, B. T. and Graesser, A. C. (2008) 'Automatic Detection of Learner's Affect from Conversational Cues', *User Modeling and User-Adapted Interaction,* 8(1-2), 45-80

Damasio, A. (1994) *Descarte's Error, Emotion, Reason and the Human Brain,* Grosset: Putnam.

Damasio, A. (2000) *The Felling of What Happens: Body and Emotion in the Making of Consciousness, Harvest Books,* 1 ed., San Diego, CA: Harcourt Brace & Co.

Darwin Correspondece Project Team (2012) 'Darwin Emotion Experiment', *Darwin Correspondence Project at University of Cambridge* [online], available: http://www.darwinproject.ac.uk/emotions [accessed 12 December 2012].

Darwin, C. R. (1872) *The Expression of the Emotions in Man and Animals,* 1 ed., London: John Murray.

De Vicente, A. (2003) *Towards Tutoring Systems that Detect Students' Motivation: An Investigation*, unpublished thesis University of Edinburgh.

De Vicente, A. and Pain, H. (1999) 'Motivation Self-report in ITS', in Lajoie, S. P. and Vivet, M., eds., *Proc. of the 9th World Conference on Artificial Intelligence in Education (AIED 99)*, Le Mans, France, 19 - 23 July 1999, Amsterdam: IOS Press, 648-650.

De Vicente, A. and Pain, H. (2002) 'Informing the Detection of the Student's Motivational State: An Empirical Study ', in Cerri, S. A., Goaurdères, G. and Paraguaçu, F., eds., *Proc. of the 6th International Conference on Intelligent Tutoring Systems (ITS 2002)*, Biarritz, France and San Sebastian, Spain, 2–7 June 2002, Heidelberg, Berlin: Springer-Verlag, 933-943.

Defino, F. (2011) 'Emotional Marketing Triggers the Right Response', *Mobile Commerce Daily: The News Leader in Mobile Commerce and Retail*, available: http://www.mobilecommercedaily.com/2011/07/28/emotional-marketing-triggers-the-right-response [accessed 28 July 2011].

Del Soldato, T. (1993) *Motivation in Tutoring Systems*, unpublished thesis The University of Sussex.

Del Soldato, T. and Du Boulay, B. (1995) 'Implementation of motivational tactics in tutoring systems', *Journal of Artificial Intelligence in Education,* 6(4), 337-378

DFKI, Paris8: Université, TUM, ICSTM, QUB and UT (2011) 'SEMAINE: The Sensitive Agent Project', *SEMAINE Project* [online], available: http://www.semaine-project.eu/ [accessed 16 January 2012].

Dias, J., Paiva, A., Vala, M., Aylett, R., Woods, S., Zoll, C. and Hall, L. (2006) 'Empathic characters in computer-based personal and social education' in Pivec, M., ed. *Affective and emotional aspects of human-computer interaction*, Netherlands, Amsterdam: IOS Press, 246-254.

Dickens, J. and Arlett, C. (2009) 'Key Aspects of Teaching and Learning in Engineering' in Fry, H., Ketteridge, S. and Marshall, S., eds., *A Handbook for Teaching and Learning in Higher Education*, Hoboken: Taylor & Francis, 264-281.

Du Boulay, B. and Luckin, R. (2001) 'Modelling Human Teaching Tactics and Strategies for Tutoring Systems', *International Journal of Artificial Intelligence in Education,* 12, 235-256.

Ekman, P. (1999) 'Basic Emotions' in Dalgleish, T. and Power, T., eds., *The handbook of Cognition and Emotion*, Sussex, UK: John Wiley and Sons, Ltd., 45-60.

Ekman, P. and Friesen, W. V. (1978) *Facial Coding System: A Technique for the Measurement of Facial Movement* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W. V. and Ancoli, S. (1980) 'Facial Signs of Emotional Experience', *Journal of Personality and Social Psychology*, 38, 270-277, available: [accessed

Elliot, A. and Pekrun, R. (2007) 'Achievement Emotion in the Hierarchical Model of Approach Avoidance Achievement Motivation' in Schutz, P. A. and Pekrun, R., eds., *Emotion in Education*, London: Elsevier, 57-74.

Er, N. F. and Dag, H. (2009) 'Comparison of Cost-free Computational Tools for Teaching Physics', in *Fifth International Conference on Soft Computing, Computing with Words and Perceptions on Systems Analysis, Decision and Control (ICSCCW 2009)*, Famagusta, North Cyprus, 2-4 September 2009, IEEE Press, 1-4.

European Comission (2008) 'LeAM: LeActiveMath consortium: language-enhanced, user adaptive, interactive elearning for mathematics', *Le Active Math: Intro* [online], available: http://www.leactivemath.org/index.html [accessed

Falcke, M. (2012) 'Relation of the group Mathematical Cell Physiology to Berlin Institute for Medical Systems Biology', *Max Delbruck Center for Molecular Medicine* [online], available: http://www.mdc-berlin.de/en/bimsb/MDC_groups/Falcke/index.html [accessed 27 July 2012].

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996) 'The KDD Process for Extracting Useful Knowledge from Volumes of Data', *Communications of the ACM,* 39(11).

Feidakis, M., Daradoumis, T. and Caballé, S. (2011) 'Emotion Measurement in Intelligent Tutoring Systems: What, When and How to Measure', in *Proc. of the 3rd IEEE International Conference on Intelligent Networking and Collaborative Systems (INCOS 2011): International Workshop on Adaptive Learning via Interactive, Collaborative and Emotional Approaches (ALICE 2011)*, Fukuoka, Japan 30 November - 2 December 2011, USA: IEEE Computer Society 807-812.

Feldman Barret, L. and Russell, J. A. (1998) 'Independence and Bipolarity in the Structure of Current Affect', *Journal of Personality and Social Psychology,* 74(4), 967-984.

Fieldman Barret, L. and Russell, J. A. (1998) 'Independence and Bipolarity in the Structure of Current Affect', *Journal of Personality and Social Psychology,* 74, 967-984.

Fowler, M. J. (2008) *Patterns of Enterprise Application Architecture,* Mass, Boston: Addison-Wesley Professional.

Francis, L., Brown, L. and Fhilipchalk, R. (1992) 'The Development of an Abreviated Form of the Revised Eysenck Personality Questionnaire (EPQR-A)', *Personality and Individual Differences,* 13, 443-449.

Freeman, D. (2004) *Creating Emotion in Games: The Craft and Art of Emotioneering,* 2 ed., Indianapolis, IN: New Riders Publishing.

Fry, H., Ketteridge, S. and Marshall, S. (2009) *A Handbook for Teaching and Learning in Higher Education,* Hoboken: Taylor & Francis.

Gage, N. L. and Berliner, D. C. (1998) *Educational Psychology,* 6 ed., Boston: Houghton Mifflin.

Gagné, R. M., Wager, W. W. and Briggs, L. J. (1992) *Principles of Instructional Design,* 4 ed., New York: Holt, Rinehart and Winston.

Games-to-Teach Team (2003) 'Design Principles of Next-Generation Digital Gaming', *Educational Technology,* 43(5), 17-33.

Gardner, H. (1983) *Frames of Mind: The Theory of Multiple Intelligences,* New York: Basic Books.

Gartner (2011) 'IT Definitions and Glossary', *Gartner Inc.* [online], available: http://www.gartner.com/technology/it-glossary/#3_0 [accessed 4 November 2011].

Gasperi, M. (2010) 'Homemade Galvanic Skin Response Meter', *Michael Gasperi's Extreme NXT Index*, available: http://www.extremenxt.com/gsr.htm [accessed 18 November 2010].

Geeknet Inc. (2010) 'Spacescape', *SourceForge* [online], available: http://sourceforge.net/projects/spacescape/ [accessed 27 October 2010].

Georgeff, M. P., Pell, B., Pollack, M. E., Tambe, M. and Wooldridge, M. (1999) 'The Belief-Desire-Intention Model of Agency', in Müller, J. P., Rao, A. S. and Singh, M. P., eds., *Proc. of*

the 5th International Workshop on Intelligent Agents V, Agent Theories, Architectures, and Languages (ATAL 1998), Paris, France, 4-7 July 1998, London, UK: Springer-Verlag, 1-10.

Glass, D., H. (2012) 'Inference to the Best Explanation: Does it track the truth?', *Synthese,* 185(3), 411-427.

Glassner, A. (2004) *Interactive Storytelling: Techniques for 21st Century Fiction,* Natick, Massachusetts: A K Peters.

Goetz, T., Frenzel, A. C., Pekrun, R., Hall, N. C. and Lüdtke, O. (2007) 'Between- and Within-Domain Relations of Student's Academic Emotions', *Educational Psychology,* 99(4), 715-733.

Goguadze, G., Sosnovsky, S., Isotani, S. and McLaren, B. M. (2011) 'Evaluating a Bayesian Student Model of Decimal Misconceptions', in Pechenizkiy, M., Calders, T., Conati, C., Ventura, S., Romero, C. and Stamper, J., eds., *Proc. of the 4th International Conference on Educational Data Mining (EDM 11)*, Eindhoven, The Netherlands, 6 - 8 July, ACM, 301-306.

Hake, R. R. (1998) 'Interactive-engagement versus Traditional Methods: A Six-thousand Student Survey of Mechanic Test Data for Introductory Physics Courses', *Ameican Journal or Physics,* 66(1), 64-74.

Han, J. and Kamber, M. (2006) *Data Mining: Concepts and Techniques, The Morgan Kaufmann Series in Data Management Systems,* 2 ed.*,* San Francisco, CA: Elsevier.

Hegarty, S. (2011) 'Crowdsourcing Darwin's Experiment on Human Emotions', *BBC News Magazine* [online], available: http://www.bbc.co.uk/news/magazine-15600203 [accessed 12 December 2012].

Henry, B. (2011) 'AutoTutor, AutoMentor Recap', *The Ohio State University: Digital Union:Office of the Chief Information Officer* [online], available: http://digitalunion.osu.edu/2011/08/24/autotutor-automentor-recap/ [accessed 11 June 2012].

Honeycutt, B. (2011) 'Applying Bloom's Taxonomy to Game Design', *Flip IT Consulting: Complimentary Article* [online], available: http://www.flipitconsulting.com/wp-content/uploads/2011/11/Article-Blooms-Taxonomy-and-Game-Design.pdf [accessed 10 December 2012].

Höök, K. (2005) 'User-Centred Design and Evaluation of Affective Interfaces.' in Ruttkoy, Z. and Pelachaud, C., eds., *From Brows to Trust: Evaluating Embodied Conversational Agents*, Netherlands: Springer, 127-160.

Hosmer, D. W. and Lemeshow, S. (2000) *Applied Logistic Regression, Wiley Series in Probability and Statistics - Applied Probability and Statistics Section,* 2 ed.*,* New York: Wiley.

Hounsell, D. (2009) 'Evaluating Courses and Teaching' in Fry, H., Ketteridge, S. and Marshall, S., eds., *A Handbook for Teaching and Lening in Higher Education*, Hoboken: Taylor & Francis, 198-211.

Huang, J., White, R. W. and Buscher, G. (2012) 'Cursor See, User Point: Gaze and Cursor Alignment in Web Search', in Konstan, J. A., Chi, E. H. and Höök, K., eds., *Proc. of the 12th SIGCHI Conference on Human Factors in Computing Systems*, Austin, Texas, 5 - 10 May 2012, New York, NY: ACM, 1341-1350

Hugin Expert A/S (2011) 'Hugin Lite - Evaluation', *HUGIN EXPERT: The Leading Support Tool* [online], available: http://www.hugin.com/productsservices/demo/hugin-lite [accessed 19 August 2011].

Hugin Expert A/S (2012) 'Hugin API: Reference Manual version 7.6', *Documentation: API Manuals* [online], available: http://download.hugin.com/webdocs/manuals/api-manual.pdf [accessed 27 July 2012].

Huitt, W. (2009) 'Humanism and Open Education', *Educational Psychology Interactive*, available: http://www.edpsycinteractive.org/topics/affect/humed.html [accessed 25 June 2012].

IBM (2012) 'SPSS Software: Predictive Analytics Software and Solutions', *Business Analytics* [online], available: http://www-01.ibm.com/software/analytics/spss/ [accessed 29 July 2012].

Jaques, P. A. and Vicari, R. M. (2007) 'A BDI Approach to Infer Student's Emotions in an Intelligent Learning Environment', *Journal of Computers & Education,* 49(2), 360-384.

Jaques, P. A., Vicari, R. M., Pesty, S. and Martin, J.-C. (2011) 'Evaluating a Cognitive-Based Affective Student Model', in D'Mello, S., Graesser, A., Schuller, B. and Martin, J.-C., eds., *Proc. of the 4th International Conference of Affective Computing and Intelligent Interaction (ACII 2011) Part I*, Memphis, TN, USA, 9-12 October 2011, Heidelberg, Berlin: Springer-Verlag, 599-608.

Jensen, F. V. and Nielsen, T. D. (2007) *Bayesian Networks and Decision Graphs, Information, Science and Statistics,* Second ed., Berlin: Springer.

JFree.org (2011) 'JFreeChart', *JFree* [online], available: http://www.jfree.org/jfreechart/ [accessed 10 May 2011].

Johnson, A. (2010) *A Little Book of Thank Yous: Letter, Notes & Quotes,* San Francisco, CA: Conari Press.

Johnson, W. L., Rickel, J. W. and Lester, J. C. (2000) 'Animated Pedagogical agents: Face to face interaction in interactive learning environments', *International Journal of Artificial Intelligence in Education,* 11(1), 47-78.

Joint Information Systems Committee (2009) 'Definitions: Technology enhanced learning environments areas', [online], available: http://www.jisc.ac.uk/whatwedo/programmes/elearning/tele/definitions.aspx [accessed 24 October 2010].

Kaya, N., Epps, H. H. and Hall, D. (2004) 'Relationship Between Color and Emotion: A Study of College Students', *College Student Journal*, 396-405, available: [accessed 8 March 2010].

Keller, J. M. (1983) 'Motivational Design of Instruction' in Reigeluth, C. M., ed. *Instructional-Design Theories and Models: An Overview of Their Current Status*, Hillsdale, NJ: Lawrence Erlbaum Associates, 386-434.

Kelly, D. and Tangney, B. (2002) 'Incorporating Learning Characteristics into an Intelligent Tutor', in Cerri, S. A., Goaurdères, G. and Paraguaçu, F., eds., *Proc. of the 6th International Conference Intelligent Tutoring Systems (ITS 2002)*, Biarritz, France and San Sebastian, Spain, 2-7 June 2002, Springer-Verlag, 729-738.

Keltner, D. and Ekman, P. (2000) 'Facial Expression of Emotion' in Lewis, M. and Haviland-Jones, J., eds., *Handbook of Emotions*, 2 ed., New York: Guilford Publications, Inc., 236-249.

Keltner, D. and Lerner, J. S. (2010) 'Emotion' in Fiske, S. T., Gilbert, D. T., Lindzey, G. and Jongsma, A. E., eds., *Handbook of Social Psychology*, 5 ed., Hoboken, N.J.: Wiley, 317-352.

Khan Acedemy (2012) 'Khan Academy: A Free World Class Education for Anyone Anywhere', *Khan Academy* [online], available: http://www.khanacademy.org/ [accessed 24 July 2012].

Khan, I. R. (2011) 'Letter to Michael Gove - Education', *Campaign of Science & Engineering in the UK* [online], available: http://sciencecampaign.org.uk/?p=2829 [accessed 12 August 2011].

Kinnear, P. R. and Gray, C. D. (2010) *PASW Statistics 17 made simple,* East Sussex: Psychology Press.

Kolb, D. A. (1984) *Experiential Learning: Experience as the Source of Learning and Development,* Englewood Cliffs, New Jersey: Prentice Hall.

Koller, D. (1999) 'Probabilistic relational models' in Džeroski, S. and Flach, P., eds., *Proc. of the 9th International Workshop of Inductive Logic Programming (ILP-99)*, Pittsburgh, USA: Springer, 3-13.

Kort, B., Reilly, R. and Picard, R. W. (2001) 'An Affective Model of Interplay Between Emotions and Learning: Reengineering Educational Pedagogy - Building a Learning Companion', in Okamoto, T., Hartley, R., Kinshuk and Klus, J. P., eds., *Proc. of the IEEE Conference on Advanced Learning Technologies*, Madison, Wiscounsin, USA, 6 - 8 August 2001, IEEE Computer Society, 43-48.

Landis, J. R. and Koch, G. G. (1977) 'The measurement of observer agreement of categorical data', *Biometrics,* 33(1), 159-174.

Lave, J. and Wenger, E. (1991) *Situated Learning: Legitimate Peripheral Participation,* Cambridge: Cambridge University Press.

Lazzaro, N. (2004) 'Why We Play Games: Four Keys to More Emotion Without Story ', *XEO Design* [online], available: http://www.xeodesign.com/xeodesign_whyweplaygames.pdf [accessed 25 October 2011].

Learning Theories (2012) 'Index of Learning Theories and Models', *Learning-Theories.com: Knowledge Base and Webliography* [online], available: http://www.learning-theories.com/ [accessed 25 June 2012].

LeDoux, J. E. (1996) *The Emotional Brain,* New York, NY: Simon & Schuster.

Leite, I., Pereira, A., Martinho, C. and Paiva, A. (2008) 'iCat, The Affective Chess Player: The Influence of Embodiment in the Enjoyment of a Game', in Petta, P. and Martinho, C., eds., *Proc. of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS08)*, Estoril, Portugal, 12-16 May, Richland, South Carolina: International Foundation for Autonomous Agents and Multiagent Systems, 1229-1232.

Lennon, R. and Eisenberg, N. (1990) 'Gender and Age Differences in Empathy and Sympathy' in Eisenberg, N. and Strayer, J., eds., *Empathy and its Development*, New York, NY: Cambridge University Press, 195 - 217.

Lepper, M. R., Woolverton, M. and Mumme, D. L. (1993) 'Motivational Techniques of Expert Human Tutors: Lessons for the design of Computer Based Tutors' in Lajoie, S. P. and Derry, S. J., eds., *Computers as Cognitive Tools*, Mahwah, N.J.: Lawrence Erlbaum Associates, 75-105.

Linnenbrink, E. A. (2007) 'The Role of Affect in Student Learning: A Multi-dimensional Approach to Considering the Interaction of Affect Motivation and Engagement' in Schutz, P. A. and Pekrun, R., eds., *Emotion in Education*, London: Elsevier, 107-124.

Luckin, R. and Du Boulay, B. (1999) 'Ecolab: The Development and Evaluation of a Vygostskian Design Framework', *International Journal of Artificial Intelligence in Education,* 10(2), 198-220.

Ma, M., Oikonomou, A. and Jain, L., eds. (2011) *Serious Games and Edutainment Applications,* London: Springer.

Mabbot, A. and Bull, S. (2006) 'Student Preferences for Editing, Persuading and Negotiating the Open Learning Model', in Ikeda, M., Ashley, K. and Chan, T. W., eds., *Proc. of the 8th International Conference on Intelligent Tutoring Systems (ITS 2006)*, Jhongli, Taiwan, 26-30 June, Berlin Heidelberg: Springer-Verlag, 481-490.

Madsen, A. L., Lang, M., Kjærulff, U. B. and Jensen, F. V. (2003) 'The Hugin Tool for Learning Bayesian Networks', in Nielsen, T. D. and Zhang, N. L., eds., *Proc. of the 7th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU-2003)*, Aalborg, Denmark, 2-5 July, Heidelberg, Berlin: Springer-Verlag, 594-605.

Malone, T. W. (1981) 'Toward a Theory of Intrinsically Motivating Instruction', *Cognitive Science,* 6(4), 333-369.

Malone, T. W. (1984) 'Heuristics for Designing Enjoyable User Interfaces: Lessons from Computer Games', in Thomas, J. C. and Schneider, M. L., eds., *Proceedings of the Conference on Human Factors in Computing Systems*, Gaithersburgh, Maryland, Norwood, NJ: Ablex Publishing Corp, 63-68.

Malone, T. W. and Lepper, M. R. (1987) 'Making Learning Fun: A Taxonomy of Intrinsic Motivations for Learning' in Snow, R. E. and Farr, M. J., eds., *Aptitude, Learning and Instruction III: Conative and Affective Process Analyses* Hillsdale, N.J: Erlbaum, 223-253.

Martinho, C., Machado, I. and Paiva, A. (2000) 'A Cognitive Approach to Affective User Modelling', in Paiva, A., ed. *Proc. of Interantional Workshop od Affect in Interactions: Towards a New Generation of Computer Interfaces (IWAI 1999)*, Siena, Italy, 20-22 October 1999, Heidelberg, Berlin: Springer-Verlag, 64-75.

Martinho, C. and Paiva, A. (2006) 'Using anticipation to create believable behaviour', in Cohn, A., ed. *Proc. of the 21st International Conference on Artificial Intelligence and the Eighteenth Innovative Applications*, Boston, Massachusetts, 16 – 20 July, 2006, USA: AAAI Press, 175-180

Marton, F. and Saljo, R. (1984) 'Approaches to Learning' in Marton, F., Hounsell, D. and Entwistle, N., eds., *The Experiences of Learning*, Edinburgh: Scottish Academic Press, 36-55.

Matsumoto, D. (1987) 'The Role of Facial Response in the Experience of Emotion: More Methodological Problems and a Meta-analysis', *Journal of Personality and Social Psychology,* 52, 769-774.

McLachlan, G. (1992) *Discriminant analysis and statistical pattern recognition,* New York, NY: John Wiley & Sons.

McQuiggan, S. W., Mott, B. W. and Lester, J. C. (2008) 'Modeling Self-efficacy in Intelligent Tutoring Systems: An inductive approach', *User Modeling and User-Adapted Interaction,* 18(1 - 2), 81-123.

Mihajlovic, V. and Petkovic, M. (2001) *Dynamic Bayesian Networks: A State of the Art*, Report, Enschede, Netherlands: University of Twente, unpublished.

Miwa, H., Takanishi, A. and Takanobu, H. (2001) 'Experimental Study on Robot Personality for Humanoid Head Robot', in *Proceedings of the International Conference of Intelligent Robots and Systems*, Maui, Hawaii, 29 October - 3 November, 2001, USA: IEEE Press, 1183-1188.

Moore, J. L., Dickson-Deane, C. and Galyen, K. (2011) 'E-learning, Online Learning and Distance Learning Environments: Are they the same?', *The Internet and Higher Education,* 14(2), 129-135.

Mootools Software (2010) 'Polygon Cruncher', *Polygon Cruncher Overview* [online], available: http://www.mootools.com/plugins/us/polygoncruncher/ [accessed 27 October 2010].

Muñoz, K., Mc Kevitt, P., Lunney, T., Noguez, J. and Neri, L. (2012) 'An Emotional Student Model for Game-based Learning' in Griol, D., Callejas, Z. and López-Cózar, R., eds., *Technologies for Inclusive Education: Beyond Traditional Integration Approaches*, Hershey, PA: IGI Global, 175-197.

Muñoz, K., Mc Kevitt, P., Noguez, J. and Lunney, T. (2009a) 'Combining Educational Games and Virtual Learning Environments for Teaching Physics with the Olympia Architecture.', in *International Symposium on Electronic Art (ISEA 2009)*, Waterfront Hall, Belfast, August 23 - September 1,

Muñoz, K., Noguez, J., Mc Kevitt, P., Neri, L., Robledo-Rella, V. and Lunney, T. (2009b) 'Adding features of educational games for teaching Physics.', in *Proceeding of the 39th IEEE International Conference Frontiers in Education*, Hotel Hilton Palacio del Rio, San Antonio, Texas, October 18-21, USA: IEEE Press, M2E-1 - M2E-6.

Murphy, K. (1998) 'A Brief Introduction to Graphical Models and Bayesian Networks', *Graphical Models* [online], available: http://www.cs.ubc.ca/~murphyk/Bayes/bnintro.html [accessed 14 December 2012].

Neji, M. and Ben Ammar, M. (2007) 'Agent-based Collaborative Affective E-learning Framework', *The Electronic Journal of e-Learning*, 5(2), 123-134, available: http://www.ejel.org/Volume-5/v5-i2/Neji.pdf [accessed 8 May 2009].

NewTek (2012) 'LightWave', *Products* [online], available: http://www.newtek.com/lightwave.html [accessed 27 March 2012].

Noguez, J. and Sucar, E. (2005) 'A semi-open learning environment for virtual laboratories', in *4th Mexican international conference on artificial intelligence: Lectures Notes in artificial intelligence*, Monterrey, México, November 14-18, Berlin: Springer Verlag, 1185-1194.

Norman, D. A., Ortony, A. and Russell, D. M. (2003) 'Affect and Machine Design: Lessons for the Development of Autonomous Machines', *IBM Systems Journal,* 42(1), 38-44.

Nwana, H. S. (1990) 'Intelligent Tutoring Systems: An Overview', *Artificial Intelligence Review,* 4, 251-277.

Oblinger, D. G. (2004) 'The Next Generation of Educational Engagement', *Interactive Media in Education*, (8), 1-18, available: http://www-jime.open.ac.uk/2004/8/ [accessed 2 March 2010].

Ohlsson, S. (1987) 'Some Principles of Intelligent Tutoring' in Lawler, R. W. and Yazdani, M., eds., *Learning Environments and Tutoring Systems*, Norwood, New Jersey, USA: Ablex Publishing, 203-237.

ORACLE Corporation (2010) 'Why MySQL?', *MySQL* [online], available: https://www.mysql.com/why-mysql/ [accessed 19 August 2011].

Ortony, A., Clore, G. L. and Collins, A. (1990) *The Cognitive Structure of Emotions,* NY: University Press.

Paireekreng, W., Rapeepisarn, K. and Wong, K. W. (2009) 'Time-based Personalised Mobile Game Downloading' in Pan, Z., Cheok, D. A. and Müller, W., eds., *Transactions on Edutainment II, LNCS 5660*, Berlin Heidelberg: Springer-Verlag, 59-69.

Papert, S. and Harel, I., eds. (1991) *Constructionism: research reports and essays 1985 - 1990 by the Epistemology and Learning Research Group, the Media Lab, Massachusetts Institute of Technology,* Norwood, NJ: Ablex Pub. Corp.

Pavlov, I. P. (1927) *Conditioned Reflexes,* London: Routledge and Kegan Paul.

Pearn, A. (2011) 'The Face of Emotions: Darwin, Computers and Autism', *Research News of the University of Cambridge* [online], available: http://www.cam.ac.uk/research/news/the-face-of-emotions-darwin-computers-and-autism/ [accessed 12 December 2012].

Pearson, N. K., Graesser, A. C. and Group, T. T. R. (2003) 'Fourteen facts about human tutoring: food for thought for ITS developers. ', in *AI-ED 2003 Workshop Proceedings on Tutorial Dialogue Systems: With a View Toward the Classroom*, Sydney Australia, 335-344.

Peek, C. (2003) 'A Primer of Biofeedback Instrumentation' in Schwartz, M. S. and Andrasik, F., eds., *BIOFEEDBACK: A Practitioner's Guide*, 3 ed., New York, NY: The Guilford Press, 43-87.

Pekrun, R., Frenzel, A. C., Goetz, T. and Perry, R. P. (2007) 'The Control Value Theory of Achievement Emotions. An integrative Approach to Emotions in Education' in Schutz, P. A. and Pekrun, R., eds., *Emotion in Education*, London: Elsevier, 13-36.

Pekrun, R., Goetz, T. and Perry, R. P. (2005) *Achievement Emotions Questionnaire (AEQ).User's manual*, Munich: University of Munich, unpublished.

Pereira, A., Martinho, C., Leite, I. and Paiva, A. (2008) 'iCat, The Affective Chess Player: The Influence of Embodiment in the Enjoyment of a Game', in Petta, P. and Martinho, C., eds., *Proc. of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS08)*, Estoril, Portugal, 12-16 May, Richland, South Carolina: International Foundation for Autonomous Agents and Multiagent Systems, 1253-1256.

Piaget, J. (1950) *The Psychology of Intelligence,* London: Routledge and Kegan Paul.

Piaget, J. (1973) *Memory and Intelligence,* New York: BasicBooks.

Picard, R. W. (1995) *Affective Computing*, Vision and Modeling, Massachussetts, USA: Massachusetts Institute of Technology (MIT), unpublished.

Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover, T., Resnick, M., Roy, D. and Strohecker, C. (2004) 'Affective learning –A Manifesto', *BT Technology Journal,* 22(4), 253-269.

Poggi, I. and Pelachaud, C. (2000) 'Performative Facial Expressions in Animated Faces' in Cassell, J., Sullivan, J., Prevost, S. and Churchill, E., eds., *Embodied Conversational Agents*, Mass, Cambridge: MIT Press, 155-158.

Porayska-Pomsta, K., Mavrikis, M. and Pain, H. (2008) 'Diagnosing and Acting on Student Affect: The tutor's perspective', *User Modeling and User-Adapted Interaction,* 18, 125-173.

Qato - Enterprise Q&A (2011) 'Finish Setup PHP, How to Communicate with MySQL and Unity', *UnityAnswers* [online], available: http://answers.unity3d.com/questions/232424/finish-setup-php-how-to-communicate-with-mysql-and.html [accessed 27 March 2011].

Qianping, W., Wei, T. and Bo, S. (2007) 'Research and Design of Edutainment', in *First IEEE International Symposium on Information Technologies and Applications in Education (ISITAE '07)*, Kunming, 23th - 25th November, IEEE, 502-505.

Rajae-Joordens, R. J. E. (2008) 'Measuring experiences in gaming and TV applications: Investigating the added value of a multi-view auto-stereoscopic 3D display' in Westerinck, J. H. D. M., Ouwerkerk, M. and Overbeek, T. J. M., eds., *Probing Experience: From Assessment of User Emotions and Behaviour to Development of Products*, Netherlands: Springer, 77-90.

Rebolledo-Mendez, G., Du Boulay, B. and Luckin, R. (2006) 'Motivating the Learner: An Empirical Evaluation' in Ikeda, M., Ashley, K. and Chan, T. W., eds., *The 8th Intelligent Tutoring Systems International Conference* Heidelberg: Springer Berlin, 545-554.

Regian, J. W., Seidel, R. J., Schuler, J. and Radtke, P. (1996) *Functional Area Analysis of Intelligent Computer-Assisted Instruction*, Washington D.C., USA: Training and Personnel Systems Science and Technology Evaluation and Management Committee (TAPSTEM), unpublished.

Reilly, R. (2008) 'Guest Editorial Virtual Laboratories: Enhancing Deep Learning in Model-Based Knowledge Domains', *IEEE Transactions in Education,* 51(1), 1.

Robinson, P. and El Kaliouby, R. (2009) 'Computation of Emotions in Man and Machines', *Philosophical Transactions of the Royal Society Biological Sciences,* 364(1535), 3441-3447.

Rollings, A. and Adams, E. (2003) *On Game Design,* Old Tappan, NJ: Pearson.

Roos, T., Wettig, H., Grünwald, P., Myllymäki, P. and Tirri, H. (2005) 'On Discriminative Bayesian Network Classifiers and Logistic Regression', *Machine Learning,* 59(3), 267-296.

Russell, S. J. and Norvig, P. (2010) *Artificial Intelligence: A Modern Approach,* 3 ed., Upper Saddle River, New Jersey: Prentice Hall.

Russell, S. J. and Norving, P. (2003) *Artificial Intelligence: A Model Approach,* 2 ed., New Jersey, USA: Pearson Education International.

Sabourin, J., Mott, B. W. and Lester, J. C. (2011) 'Modelling Learner Affect with Theoretical Grounded Dynamic Bayesian Networks', in D'Mello, S., Graesser, A., Schuller, B. and Martin, J.-C., eds., *Proc. of the 4th International Conference of Affective Computing and Intelligent*

Interaction (ACII 11) Part I, Memphis, TN, USA, 9-12 October 2011, Heidelberg, Berlin: Springer-Verlag, 286-295.

Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan , C. and Bigdeli , A. (2008) 'How do you know that I don't understand? A look at the future of intelligent tutoring systems', *Computers in Human Behavior,* 24(4), 1342-1363.

Schaller, D. T., Allison-Bunnell, S. and Borun, M. (2005) 'Learning Styles and Online Interactives', *Proceedings of Archives and Museum Informatics, Museums and the Web 2005* [online], available: http://www.archimuse.com/mw2005/papers/schaller/schaller.html [accessed 31 March].

Schank, R. C. (1996) *Tell Me a Story, Rethinking Theory,* Evanston, Illinois: Northwestern University Press.

Schank, R. C., Fano, A., Bell, B. and Jona, M. (1994) 'The Design of Goal-Based Scenario', *The Journal of the Learning Sciences,* 3(4), 305-345.

Schutz, P. A. and Pekrun, R., eds. (2007) *Emotion in Education,* San Diego, California, USA: Elsevier.

Schwartz, M. S. and Olson, R. P. (2003) 'A Historical Perspective on the Field of Biofeedback and Applied Psychophysiology' in Schwartz, M. S. and Andrasik, F., eds., *BIOFEEDBACK: A Practitioner's Guide*, 3 ed., New York, NY: The Guilford Press, 3-19.

Serway, R. A. and Jewett, J. W. (2004) *Physics for Scientists and Engineers,* 6 ed., Brooks Cole.

Sillitto, R. and MacKinnon, L. M. (2000) 'Going Splat! - Building a Multimedia Educational Resource for Physics Learners', *Physics Education,* 35(5), 325-331.

Singhal, A. and Brown, C. (1997) 'Dynamic Bayes Net Approach to Multimodal Sensor Fusion', *Proc.of the International Society for Optical Engineering (SPIE): Sensor Fusion and Decentralized Control in Autonomous Robotic Systems*, 3209, 2-10, available: http://dx.doi.org/10.1117/12.287628 [accessed 22nd September].

Skinner, B. F. (1950) 'Are Theories of Learning Necessary?', *Psychological Review,* 57, 193-216.

Smith, L. B. and Breazeal, C. (2007) 'The Dynamic Lift of the Developmental Process', *Developmental Science,* 10(1), 61-68.

Smith Micro Inc. (2010) 'Poser 8', *Composer: Complete 3D Figure Design & Animation* [online], available: http://poser.smithmicro.com/poser.html [accessed 12 December 2010].

Sound-effects-library (2010) 'Sound-effects-library: The World's Largest Audiotionable Sound Effects, Music Samples and Music Tracks Library', *Sound-effects-library* [online], available: http://sound-effects-library.com/index.asp [accessed 18 December 2010].

SourceForge (2008) 'BlueCove', *SourceForge.net: BlueCove* [online], available: http://bluecove.org/ [accessed 18 December 2010].

SourceForge (2009) 'LeJOS: Java for Lego Mindstorms', *SourceForge.net: LeJOS* [online], available: http://lejos.sourceforge.net/index.php [accessed 18 December 2010].

Souto, M. A. M., Nicolao, M., Vicari, R. M., de Oliveira, J. P. M., Verdin, R., Beschoren, K., Madeira, M. and Zanella, R. (2002) 'Web-Adaptive Training System Based on Cognitive Student Style', in Passey, D. and Kendall, M., eds., *Proc. of International Federation for Information Processing (IFIP 2002) 17th World Computer Congress--TC3 stream on TelE-LEARNING: The Challenge for the Third Millenium*, Montréal, Québec, Canada, 25-30 August 2002, Norwell, Massachusetts: Kluwer Academic Publishers, 281-288.

Squire, K. (2006) 'From Content to Context: Video Games As Designed Experiences', *Educational Researcher,* 35(8), 19-29.

Starostin, P. (2007) 'LowMax', *LowMax - Free Character Rigs Downloads for 3dsmax* [online], available: http://www.creativecrash.com/3dsmax/downloads/character-rigs/c/lowmax [accessed 15 June 2010].

Sucar, L. E. and Noguez, J. (2008) 'Student Modeling' in Pourret, O., Naïm, P. and Marcot, B., eds., *Bayesian Networks: A Practical Guide to Applications* West Sussex, England: J. Wiley & Sons, 173-185.

Sykes, J. (2006) 'Affective Gaming: Advancing the Argument for Game-Based Learning' in Pivec, M., ed. *Affective and Emotional Aspects of Human-Computer Interaction* Netherlands, Amsterdam: IOS Press, 3-7.

Tabacknick, B. G. and Fidell, L. S. (2007) *Using Multivariate Statistics,* 5 ed., Boston, MA: Pearson.

TechSmith (2011) 'Camtasia: Screen Recording Software', *Camtasia 2* [online], available: http://www.techsmith.com/camtasia2-1111.html?gclid=CIL3hcfhmLACFRQrfAod6imP1w [accessed 15 August 2011].

The Apache Software Foundation (2011) 'Apache Tomcat', *The Apache Jakarta Project* [online], available: http://tomcat.apache.org/ [accessed 19 August 2011].

The Games-to-Teach Research Team (2003) 'Design Principles of Next-Generation Digital Gaming for Education', *Educational Technology,* 43(5), 17-33.

The ICE House Project Group (2009) '*Teaching for Innovation, Creativity and Enterprise*', [online], available: http://www.ice-house.info/Resources/ICE_House_Project_1st_Briefing_Paper.pdf [accessed 2 March 2010].

The University of Waikato (2012) 'WEKA 3: Data Mining Software in Java', *WEKA: The University of Waikato* [online], available: http://www.cs.waikato.ac.nz/ml/weka/ [accessed 20 July 2012].

Thrun, S. (2002) 'Probabilistic Robotics', *Communications of the ACM,* 45(3), 52-57.

Tu Graz, COVUNI, UOC, MOMA and CRMPA (2012) 'ALICE: Adaptive Learning via Intuitive/Interactive Collaborative and Emotional Systems', *ALICE Project* [online], available: http://www.aliceproject.eu/ [accessed

TurboSquid (2010) 'TuboSquid: 3D Models, 3D Modeling Textures and Plugins', *TurboSquid* [online], available: http://www.turbosquid.com/ [accessed 11 November 2010].

Turkle, S., Breazeal, C., Dasté, O. and Scassellati, B. (2006) 'First Encounters with Kismet and Cog: Children Respond to Rationale Artifacts' in Messaris, P. and Humphrey, L., eds., *Digital Media: Transformations in Human Communication*, New York, NY: Peter Lang Publishing, Inc., 313-330.

Udacity Inc (2012) 'Udacity: Twenty-first Century University', *Udacity* [online], available: http://www.udacity.com/ [accessed 24 July 2012].

Unity Technologies (2011) 'Unity 3', [online], available: http://unity3d.com/unity/ [accessed 19 August 2011].

Van Eck, R. (2006) 'Digital Game-Based Learning: It's Not Just the Digital Natives Who Are Restless', *EDUCAUSE Review*, 41(2), 16-30, available: http://www.educause.edu/EDUCAUSE+Review/EDUCAUSEReviewMagazineVolume41/DigitalGameBasedLearningItsNot/158041 [accessed 24 October 2011].

Virvou, M., Manos, K. and Katsionis, G. (2006) 'Virtual Reality Edutainment: Cost Effective Development of Personalised Software Applications', in Gervasi, O., Kumar, V., Tan, C. J. K., Taniar, D., Laganà, A., Mun, Y. and Choo, H., eds., *Proc.of Computational Science and its Applications (ICCSA 2006), Part I*, Glasgow, UK, 8 -11 May 2006, Springer, 251-260.

Visual Paradigm (2011) 'Visual Paradigm: Boost Productivity with Innovative and Intuitive Technologies', *Visual Paradigm Suite* [online], available: http://www.visual-paradigm.com/product/?favor=bpva [accessed 9 November 2011].

Vygotsky, L. S. (1998) *Formação Social da Mente o Desenvolvimento dos Processos Psicológicos Superiores,* 6 ed., São Paulo, Brazil: Editora Martins Fontes, Ltd.

Westerinck, J. H. D. M. (2008) 'Experience in Products' in Westerinck, J. H. D. M., Ouwerkerk, M. and Overbeek, T. J. M., eds., *Probing Experience: From Assessment of User Emotions and Behaviour to Development of Products*, Netherlands: Springer, 1-8.

Westerinck, J. H. D. M., Ouwerkerk, M., Overbeek, T. J. M., Pasveer, W. F. and De Ruyter, B. (2008a) *Probing Experience: From Assessment of User Emotions and Behaviour to Development of Products,* Dordrecht, The Netherlands: Springer.

Westerinck, J. H. D. M., Van den Broek, E. L., Schut, M. H., Van Herk, J. and Tuinenbreijer, K. (2008b) 'Computing emotion awareness though galvanic skin response and facial electromiography' in Westerinck, J. H. D. M., Ouwerkerk, M. and Overbeek, T. J. M., eds., *Probing Experience: From Assessment of User Emotions and Behaviour to Development of Products*, Netherlands: Springer, 149-162.

Wilson, D. (2010) 'Google Nabs Patent to Monitor Your Cursor Movements', *TechEye.net*, available: http://news.techeye.net/internet/google-nabs-patent-to-monitor-your-cursor-movements [accessed 11 November 2010].

Witten, I. H. and Frank, E. (2005) *Data Mining: Practical Machine Learning Tools and Techniques, Series in Data Management Systems,* San Francisco, CA: Elsevier Inc.

Wittgenstein, L. (1953/2009) *Philosophical Investigations,* 4 ed., Chichester, West Sussex: Wiley-Blackwell.

Wooldridge, M. (2002) *An Introduction to Multi-Agent Systems,* West-Sussex, England: John Wiley & Sons Ltd.

Woolf, B. P. (2009) *Building Intelligent Interactive Tutors: Student-centered strategies for revolutionizing e-learning,* Burlington, NJ, USA: Elsevier, Inc.

Zammito, V. L. (2005) 'The Expression of Colours', in Castell, S. D. and Jenson, J., eds., *Proceedings of the 2005 Digital Games Research Association Conference, Changing Views: Worlds in Play*, Vancouver, 16 - 20 June, University of Vancouver, 15.