3rd Workshop on Personalization Approaches for Learning Environments (PALE 2013)

Preface

Milos Kravcik[1], Olga C. Santos[2], Jesus G. Boticario[2], Diana Pérez-Marín[3]

[1] RWTH University Aachen, Germany
kravcik@dbis.rwth-aachen.de
[2] aDeNu Research Group, Artificial Intelligence Department,
Computer Science School, UNED, Spain
ocsantos@dia.uned.es – jgb@dia.uned.es
http://adenu.ia.uned.es/
[3] Laboratory of Information Technologies in Education (LITE).
Universidad Rey Juan Carlos, Spain
diana.perez@urjc.es

Abstract. Personalization approaches in learning environments can be addressed from different perspectives and also in various educational settings, including formal, informal, workplace, lifelong, mobile, contextualized, and self-regulated learning. PALE workshop offers an opportunity to present and discuss a wide spectrum of issues and solutions, such as pedagogic conversational agents, personal learning environments, and learner modeling.

1 Introduction

The 3rd International Workshop on Personalization Approaches in Learning Environments (PALE)[1] takes place on June 10th, 2013 and is held in conjunction with the 21th conference on User Modeling, Adaptation, and Personalization (UMAP 2013). The topic can be addressed from different and complementary perspectives. PALE workshop aims to offer a fruitful crossroad where interrelated issues can be contrasted, such as pedagogic conversational agents, responsive open learning environments, and learner modeling. The benefits of the personalization and adaptation of computer applications have been widely reported both in e-learning (the use of electronic media to teach, assess, or otherwise support learning) and b-learning (to combine traditional face-to-face instruction with electronic media - blended learning).

[1] http://adenu.ia.uned.es/workshops/pale2013/

Previous PALE workshops (both at UMAP 2011 and UMAP 2012) have shown several important issues in this field, such as behavior and embodiment of pedagogic agents, suitable support of self-regulated learning, appropriate balance between learner control and expert guidance, design of personal learning environments, contextual recommendations at various levels of the learning process, predicting student outcomes from unstructured data, modeling affective state and learner motivation, and using sensors to understand student behavior and tracking affective states of learners, harmonization of educational and technological standards, processing big data for learning purposes, predicting student outcomes, adaptive learning assessment, and evaluation of personalized learning solutions. This points at individualization of learning as still a major challenge in education where rapid technological development brings new opportunities how to address it. A lot of data can be collected in the educational process, but we need to find ways how to use it reasonably and to develop useful services in order to make the learning process more effective and efficient. Novel personalized services and environments are needed especially in lifelong and workplace educational settings, in order to support informal, self-regulated, mobile, and contextualized learning scenarios. A big challenge is also adaptation considering both long-term objectives and short-term dynamically changing preferences of learners. Here open and inspectable learner models play an important role. In the case of pedagogic conversational agents personalization is fostered by the use of adapted dialogues to the specific needs and level of knowledge of each student.

In order to foster the sharing of knowledge and ideas to research on these issues, PALE format moves away from the classic 'mini-conferences' approach and follows the Learning Cafe methodology to promote discussions on open issues regarding personalization in learning environments. This means that participants attending the workshop benefit both from interactive presentations and constructive work.

2 Workshop themes

The higher-level research question addressed in the workshop is: "What are suitable approaches to personalize learning environments?" It is considered in various contexts of interactive, personal, and inclusive learning environments. The topics of the workshop included (but not limited to) the following:

- Motivation, benefits, and issues of personalization in learning environments
- Approaches for personalization of inclusive, personal and interactive learning environments
- Successful methods and techniques for personalization of learning environments
- Results and metrics in personalized learning environments
- Social and educational issues in personalized learning environments
- Use of pedagogic conversational agents
- Affective computing in personalized learning environments
- Ambient intelligence in personalized learning environments
- User and context awareness in personalized learning environments

3 Contributions

A blind peer-reviewed process by three reviewers per paper with expertise in the area was carried out to select the contributions for the workshop. As a result, 4 submissions were accepted, which report designing approaches, evaluation methods and open issues for eliciting the recommendation support to personalize learning environments.

Arevalillo-Herráez et al. [1] discuss what is needed to design an experiment for capturing relevant information from an ITS to improve the learner's competence in solving algebraic word problems considering learners' emotional and mental states. To enrich learner's experience with affective support both action logs to record user's interaction with the system, which can be used to discover important information that help instructional designers to improve the ITS performance, and emotional information gathered from external sources, which reflect affective or mental states, can be used.

Labaj and Bieliková [2] propose a conversational evaluation approach be used within ALEF adaptive learning framework that tracks the user attention and uses that information to ask the evaluation questions at the appropriate time and right when the user is working with the part in question (or just finished working with it). This approach aimed to get higher cooperation from the user providing more feedback than when we would ask them randomly.

Koch et al. [3] are researching, developing, and testing technologies to instrument classrooms, collect human signal data, and derive meaning that can lead to understand their relation with the education performance. In particular, they have developed an interface to capture human signals in learning environment, integrated into innovative analytic models to extract meaning from these data and have implemented a proof-of-concept experiment to detect variations of attention deficit hyperactivity disorder based on level of attentiveness, activity and task performance.

Manjarrés-Riesco et al. [4] discuss open issues which arise when eliciting personalized affective recommendations for distance learning scenarios, such as scarce reported experiences on affective support scenarios, ii) affective needs, iii) difficulties of affective communication in virtual learning communities, iv) reduced scope of the affective support provided in current approaches, and v) lack of resources for educators to provide affective support. These issues were identified in the course of applying TORMES user centered engineering approach to involve relevant stakeholders (i.e. educators) in an affective recommendation elicitation process.

Acknowledgements

PALE chairs would like to thank the authors for their submissions and the UMAP workshop chairs for their advice and guidance during the PALE workshop organization. Moreover, we also would like to thank the following members of the Program Committee for their reviews: Miguel Arevalillo, Maria Bielikova, Zoraida Callejas, Cristina Conati, Sabine Graf, David Griol, Judy Kay, Kinshuk, Ralf Klamma, Tobias

Ley, Ramón López-Cózar, Noboru Matsuda, Beatriz Mencía, Alexander Nussbaumer, Alexandros Paramythis, Dimitris Spiliotopoulos, Carsten Ullrich, and Martin Wolpers. The organization of the PALE workshop relates and has been partially supported by the following projects: ROLE (FP7 IST-231396), Learning Layers (FP7 318209) funded by the 7FP of the European Commission, and MAMIPEC (TIN2011-29221-C03-01) funded by the Spanish Ministry of Economy and Competence.

References

1. Arevalillo-Herráez, M., Moreno-Picot, S., Arnau, D., Moreno-Clari, P., Boticario, J.G., Santos, O.C., Cabestrero, R., Quirós, P., Salmeron-Majadas, S., Manjarrés-Riesco, A., Saneiro. Towards Enriching an ITS with Affective Support. In proceedings of the 3nd Workshop on Personalization Approaches for Learning Environments (PALE 2013). Kravcik, M., Santos, O.C., Boticario, J.G. and Pérez-Marín, D. (Eds.). 21th conference on User Modeling, Adaptation, and Personalization (UMAP 2013), 2012, p. 5-13.
2. Labaj, M. Bieliková, M. Conversational Evaluation of Personalized Solutions for Adaptive Educational Systems. In proceedings of the 3nd Workshop on Personalization Approaches for Learning Environments (PALE 2013). Kravcik, M., Santos, O.C., Boticario, J.G. and Pérez-Marín, D. (Eds.). 21th conference on User Modeling, Adaptation, and Personalization (UMAP 2013), 2012, p. 14-19.
3. Koch, F., Ito, M., da Silva, A.B.M., Borger, S., Nogima, J. Exploiting Human Signals in Learning Environment as an Alternative to Evaluate Education Performance. In proceedings of the 3nd Workshop on Personalization Approaches for Learning Environments (PALE 2013). Kravcik, M., Santos, O.C., Boticario, J.G. and Pérez-Marín, D. (Eds.). 21th conference on User Modeling, Adaptation, and Personalization (UMAP 2013), 2012, p. 20-25.
4. Manjarrés-Riesco, A., Santos, o.C., Boticario, J.G., Saneiro, M. Open Issues in Educational Affective Recommendations for Distance Learning Scenarios. In proceedings of the 3nd Workshop on Personalization Approaches for Learning Environments (PALE 2013). Kravcik, M., Santos, O.C., Boticario, J.G. and Pérez-Marín, D. (Eds.). 21th conference on User Modeling, Adaptation, and Personalization (UMAP 2013), 2012, p. 26-33.

Towards Enriching an ITS with Affective Support

Miguel Arevalillo-Herráez[1], Salvador Moreno-Picot[1], David Arnau[2], Paloma Moreno-Clari[1], Jesus G. Boticario[3], Olga C. Santos[3], Raúl Cabestrero[4], Pilar Quirós[4], Sergio Salmeron-Majadas[3], Ángeles Manjarrés-Riesco[3], Mar Saneiro[3]

[1] Department of Computer Science, University of Valencia, Spain
miguel.arevalillo@uv.es, salvador.moreno@uv.es, paloma.moreno@uv.es
[2] Department of Mathematics Education, University of Valencia, Spain
david.arnau@uv.es
[3] aDeNu Research Group. Artificial Intelligence Dept. Computer Science School. UNED, Spain
jgb@dia.uned.es, ocsantos@dia.uned.es, ssalmeron@bec.uned.es, amanja@dia.uned.es, marsaneiro@dia.uned.es
[4] Department of Basic Psychology, Faculty of Psychology, UNED, Spain
rcabestrero@psi.uned.es, pquiros@psi.uned.es

Abstract. Recent progress in affective computing is having an important impact on the development of Intelligent Tutoring Systems (ITS). Many ITS use action logs to record user's interaction with the system, such as to discover important information that help instructional designers to improve the ITS performance. However, finer grain interaction data as well as emotional information gathered from external sources is required to determine affective or mental states that can be used to enrich learner's experience with affective support. In this paper, we discuss what is needed to design an experiment for capturing relevant information from an ITS to improve the learner's competence in solving algebraic word problems considering learners' emotional and mental states.

Keywords: Affective computing, ITS, Multimodal emotions detection.

1 Introduction

User's affective state features a strong relationship with the cognitive process [1-4]. In the MAMIPEC and MARES projects we aim at exploring potential applications of affective computing in the context of accessible and personalized learning systems. To this end, we consider a user context that includes a wide range of appliances and devices to enrich the user's interaction. To study possible ways to detect user's emotions in a learning context, a number of experiments focused on emotional data gathering have been carried out. A total of 92 subjects with different profiles and backgrounds, including people with functional diversity [5], were asked to solve a collection of mathematical exercises through dotLRN Learning Management System (LMS) while emotional information was gathered both from sensors and questionnaires.

In order to further understand the learning implications of affective states, identify possible applications of affect detection in tutoring systems, and reinforce some of the conclusions drawn from the above study, we are currently following two research directions: 1) investigating potential applications of affective computing to improve an ITS developed in the context of the MARES project [6, 7]; 2) extending the dotLRN open source LMS and related software modules to include the required adaptive affective support through affective educational oriented recommendations [8].

This paper describes some of the actions adopted by both research groups to improve the existing ITS and endow it with adaptive and affective support through recommendations. This ITS is deployed as a standalone application that provides tutoring features on a mathematical topic. In particular, the application aims at improving the learner's competence in solving algebraic word problems. The algebra domain has been chosen because of the many possibilities that it offers, regarding potential responses to specific mental states. Next, the ITS is described. After that, we discuss how to enrich the ITS with affective information based on the analysis of results carried out to date on the aforementioned experiments.

2 ITS description and position within the state of the art

The ITS emulates the behavior of a human tutor by tracking the current resolution path that the student is following, and adapts feedback accordingly. To this end, expert knowledge on the structure of word problems is codified by using hypergraphs that represent the relations between quantities in the different analytical readings associated with each problem [6, 7]. The system is able to provide feedback and hints on demand. In both cases, the most likely analytical reading is computed and used to adapt the system response, which is given in natural language.

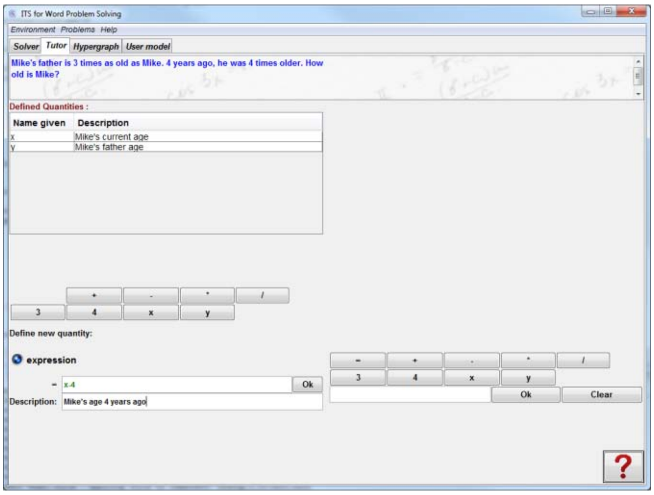

Fig. 1. A screenshot of the ITS in tutoring mode

Fig. 1 shows a screenshot of the system in tutoring mode. The panel on the left hand side is used to define quantities, either by using a letter or as a function of other quantities that have already been defined. In the figure, the student has already used letters x and y to designate the ages of Mike and his father, respectively; and is currently defining Mike's age 4 years ago as x-4. The panel on the right hand size is used to build equations that relate several existing quantities. To encourage a systematic problem solving approach, calculator-like components are used in both cases. These contain the basic operators and one button per quantity already defined. The component used to build equations includes an additional button for the equals sign. In this way, quantities need to be defined before they are used to either define another quantity or set an equation. The question mark button at the right-bottom corner of the screen is used to request a hint. If this button is pressed, a hint is displayed on a floating window. This window is also used to provide feedback to incorrect actions. In Fig.1, a sample help box is also shown on top of the main application window.

The ITS has been designed so that action logs are dynamically produced as the user interacts with the system. Student actions are written to a file in natural language. Fig. 2 shows an example of the output generated. In this file, it can be observed that after defining the two letters, the student requested a hint. Again, the student felt unable to carry out the recommended action and asked for further help. The system reacted by giving further details on the first action suggested. Still, the student did not know what to do and abandoned the application without finishing the resolution. Apart from other obvious uses of such visual information (e.g. files can be inspected to study the student's performance in detail), we are currently working on applying machine learning algorithms to the logs in order to draw relevant conclusions regarding situations that may demotivate the learner and cause abandonments.

```
NEW PROBLEM LOADED: Ages
    STATEMENT: Mike's father is 3 times as old as Mike. 4 years ago, he
was 4 times older. How old is Mike?
    USER ACTION: DEFINING LETTER.
      - x to represent Mike's current age
    SYSTEM ACTION: ACCEPTED
    USER ACTION: DEFINING LETTER.
      - y to represent Mike's father age
    SYSTEM ACTION: ACCEPTED
    SYSTEM ACTION: HINT GIVEN.
      4 years ago, Mike was four years younger than today
      You may try to define
        Mike's age 4 years ago
        as a function of
          - 4
          - Mike's current age (x)
    SYSTEM ACTION: HINT GIVEN.
      4 years ago, Mike was four years younger than today
      Hence
        Mike's age 4 years ago = Mike's current age less 4
        You may try to define
        Mike's age 4 years ago
        as x-4
    USER ACTION: EXITING WITHOUT FINISHING
```

Fig. 2. An example of the high-level log produced by the application

Despite the possibilities offered by the high level information in the logs, finer grain interaction data may have a relatively higher importance to determine affect or mental states. For example, inactivity times, mouse movements or the time elapsed between clicks when defining an expression may provide important indicators relevant for the learning process. Combined with other (ideally non-invasive) sources of information (webcams, eye tracking hardware), interaction data can be used to detect specific emotional situations such as concentration, boredom, confusion or frustration [9-11]. In turn, this information can be directly used by the ITS to adapt common responses and/or handed to a recommender system to act in consequence [2].

3 Issues to consider for emotions detection in the ITS

Currently, we are trying to take advantage of the ITS tracking capabilities to enrich the multimodal emotional data mining detection approach [12] by gathering more detailed interaction and emotional information from the ITS and further exploit its adaptive features. In a previous experience [5], which was planned by a multidisciplinary team that includes experts in different fields (mathematics education, psychology, programming, data mining, machine learning and modeling), participants had to solve multiple choice mathematical exercises. Affective states were elicited at predetermined moments during the experience and gathered through several sources as follows: i) physiological sensors (hearth rate, breath rate, temperature, galvanic skin response, blood pressure) in order to detect significant variations related to certain changes in learner's affective state, ii) video recording (web cams, Kinect device, eye tracker) to find characteristic emotional meaningful facial gestures and attention foci, iii) interaction records (from mouse, keyboard and desktop) to identify behavioural changes, iv) standardized questionnaires (e.g. Big Five Inventory, General Self-Efficacy Scale, Positive and Negative Affect Schedule) to take into account certain aspects of participant’s personality and emotions and v) self-reports and scales (e.g. Self-Assessment Manikin) on their feelings and thoughts.

In order to elicit several affective states, three groups of questions were prepared. The first one was easy if paper and pencil could be used, but participants were not allowed to do so. The next group of questions was limited in time, allowing less time than needed in order to cause stress in the participants (they were told that time was sufficient enough to fulfill the task). In the third group of questions the difficulty level was lowered and the type of problem was changed to logical series in order that participants could finish the session with a sensation of joy and happiness.

Although the experimental design allowed for a coherent data capture, the platform (being an LMS) lacked of some relevant functionality that could have enriched the quality of the information gathered in order to dynamically adapt the system behavior according to the user’s input. Moreover, dotLRN does not provide in-built support for capturing interaction data. Mouse clicks, keyboard strokes and other interaction data had to be captured by using independent software, hence requiring a careful post-processing step to ensure that all data were adequately synchronized with the additional sensors used, namely video recorded information and physiological sensors.

In order to plan the experimental design to be carried out with the ITS based on the experience of previous experiments, the following issues are being addressed:

- **Determination of the affective and mental states that are relevant from a learning perspective.** Boredom, interest and frustration are some of these relevant states. Here, the involvement of educators with experience in providing affective support through virtual learning scenarios is needed. The purpose here is to help these educators to identify relevant situations that require emotional support, and ultimately define the support to be provided. TORMES methodology, which adapts the ISO standard 9241-210, can be used to guide educators to elicit and describe affective recommendations with educational value in their scenarios [13].
- **Selection of the most adequate devices to capture appropriate data that can be used to infer the user's affective/mental state**. Although non-intrusive devices are preferred, other more intrusive are still of interest. Despite that they may not be directly applicable in current practical setups, research conclusions may highlight their importance and encourage the construction of non-intrusive devices to capture the same type of signal. Microsoft Kinect technology, eye tracking, webcams, physiological sensors were already used at previous experiments [5, 12].
- **Plan the data gathering process.** This includes deciding on the most relevant variables, the format used to record the data produced by the different devices, and the synchronization mechanism that will be used to be able to combine information coming from the multiple sources. The ITS runs as a standalone application, and may easily be extended to capture low level interaction data related to keystrokes and mouse clicks and movements, and/or modified to adapt the higher level information that is currently recorded to the objectives of the experience. However, other devices are not integrated into the ITS and may require further development to ease the subsequent analysis, as well as their synchronization.
- **Elicitation of affective/mental states.** During the interaction, affective and mental states need to be provoked. The analysis of this type of interventions is focused on inferring changes in the user's affective state from the reactions detected by the input devices. The ITS is currently able to assess the difficulty of each problem by examining the analytical reading associated with them; and it is also capable of providing adapted feedback by using parameterized templates. These two features can be exploited to devise specific instructional designs aimed at eliciting emotions rather than maximizing learning.

4 Major Challenges

Apart from the intrinsic difficulty associated with detecting metal states in a non-intrusive way, there are many other aspects that make the work on enriching the ITS with affective support specially challenging. Currently, the ITS does not incorporate a recommender system, which could issue appropriate responses when particular affective and mental states are detected. The integration of the many aspects involved in the construction of a recommender system based on the detection of mental states requires expertise in several application fields. For this reason, a multidisciplinary team with experience in different areas has been built. Group members coming from the psychology field have previously copped with the problems involved in the affective states detection. In particular, their experience will serve to collect and integrate the great variety of sources of information –cognitive, behavioral and physiological– present in the study of emotions. This will be specifically valuable to maximize the amount, accuracy and relevance of emotional information, along with the minimization of intrusiveness to yield more accurate information. The expertise of other members in artificial intelligence is needed to construct data models which are appropriate for the problem at hand, to design the inference that will support the system and to combine the multiple information sources which will be fed into the recommender. Psycho-educational expertise is another fundamental ingredient, mainly related to the identification of situations where recommendations may have a positive impact in learning. Some team members have extensive experience in this subject and have developed an entire methodology to support the elicitation of educational oriented recommendations (TORMES), which can be used to identify opportunities where affective based recommendations could be offered [13]. Emotional support in e-learning platforms is currently a widely addressed issue in order to take advantage of the role emotions play in learning and cognitive process [1]. Affective state detection is a necessary open issue widely addressed, where the use of many different data sources (drawing a multimodal approach) is being applied in order to get new and richer information about the learner [14]. Due to the huge amounts of data a multimodal approach can lead to, the use of data mining techniques to extract affective information from the data gathered surfaces.

5 Conclusions and Future Work

Results from first experience on emotional states detection is the basis for a new experiment that aims to incorporate affective support to an existing ITS in the algebra domain. Shortcomings identified in the first experience have been considered to be used in a more flexible standalone tutoring application (from the adaptation point of view) than an LMS. Experiments are being designed on the same platform as affective support will be provided (i.e. the ITS).

Once the new experiment is run, data mining will proceed in a similar way as in the previous experiments. Conclusions will be used to include affective information into the user's model, and to adapt the ITS to react to disruptive emotional and mental states. The effect of incorporating affective support will then be evaluated in a real environment, with the participation of students at secondary education. High level logs will also be analyzed to identify factors that may contribute to promoting positive and negative mental states. This information will be used by instructional designers to improve the ITS. Furthermore, it can also be used to define affective educational oriented recommendations that can deliver affective feedback provided that the required adaptive infrastructure has been developed. This project implies far more than simply detecting mental states, but the development of a module that uses information related to the user's mental states to improve learning is a challenging issue on itself.

Acknowledgments

Authors would like to thank the Spanish Ministry of Economy and Competitivity for funding the MAMIPEC and MARES projects (TIN2011-29221-C03-01 and TIN2011-29221-C03-02), which have supported this research work.

References

1. Porayska-Pomsta, K., Mavrikis, M., and Pain, H., "Diagnosing and acting on student affect: the tutor's perspective," *User Modeling and User-Adapted Interaction*, vol. 18, no. 1–2, pp. 125–173, Feb. 2008.
2. Shen, L., Wang, M., and Shen, R., "Affective e-Learning: Using 'Emotional' Data to Improve Learning in Pervasive Learning Environment," *Educational Technology & Society*, vol. 12, no. 2, pp. 176–189, 2009.
3. Blanchard, E. G. , Volfson, B., Hong, Y.-J. and Lajoie, S. P. ,"Affective Artificial Intelligence in Education: From Detection to Adaptation," in *Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling*, Amsterdam, The Netherlands, The Netherlands, 2009, pp. 81–88.
4. Forgas, J. P. , Baumeister, R. F., and Tice, D. M. *Psychology of self-regulation: Cognitive, affective, and motivational processes.*, vol. 11. New York, NY US: Psychology Press, 2009.
5. Santos, O. C., Rodríguez-Ascaso, A., Boticario, J. G., Salmeron-Majadas, S., Quirós, P., Cabestrero, R. Challenges for inclusive affective detection in educational scenarios. In proceedings of HCI International 2013. Lecture Notes in Computer Science, in press (2013)
6. Arevalillo-Herráez, M., Arnau, D., Marco-Giménez, L., Domain-specific knowledge representation and inference engine for an intelligent tutoring system, Knowledge Based Systems (2013), in press.
7. Arnau, D., Arevalillo-Herráez, M., Puig, L., González-Calero, J.A., Fundamentals of the design and the operation of an intelligent tutoring system for the learning of the arithmetical and algebraic way of solving word problems. Computers & Education 63: 119-130 (2013)
8. Santos, O.C., Boticario, J.G., Manjarrés-Riesco, A. An approach for an Affective Educational Recommendation Model. Recommender Systems for Technology Enhanced Learning: Research Trends & Applications. Manouselis, N., Drachsler, H., Verbert, K., Santos, O.C. (eds.). Springer, 2013 (in press).
9. Elliott, C., Rickel, L., Lester, J.. 1999. Lifelike pedagogical agents and affective computing: an exploratory synthesis. In Artificial intelligence today, Michael J. Wooldridge and Manuela Veloso (Eds.). Lecture Notes In Computer Science, Vol. 1600. Springer-Verlag, Berlin, Heidelberg 195-212.
10. de Vicente, A., Pain, H., "Motivation Diagnosis in Intelligent Tutoring Systems" Intelligent Tutoring Systems. Lecture Notes in Computer Science Volume 1452, 1998, pp 86-95
11. D'Mello, S.; Picard, R.W.; Graesser, Arthur, "Toward an Affect-Sensitive AutoTutor," Intelligent Systems, IEEE , vol.22, no.4, pp.53,61, July-Aug. 2007

12. Santos, O.C., Salmeron-Majadas, S., Boticario, J.G. Emotions detection from math exercises by combining several data sources. In proceedings of the 16th International Conference on Artificial Intelligence in Education (AIED 2013), in press
13. Santos, O.C., Boticario, J.G. (2012) Affective Issues in Semantic Educational Recommender Systems. In Proceedings of the 2nd Workshop on Recommender Systems for Technology Enhanced Learning. CEUR-WS proceedings, vol. 896, p. 71-82.
14. D'Mello, S. K., & Kory, J. Consistent but Modest: A Meta-Analysis on Unimodal and Multimodal Affect Detection Accuracies from 30 Studies. In L. P. Morency et al (Eds)Proceedings of the 14th ACM International Conference on Multimodal Interaction (2012). (pp. 31-38). New York:ACM

Conversational Evaluation of Personalized Solutions for Adaptive Educational Systems

Martin Labaj, Mária Bieliková

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova, 842 16 Bratislava, Slovakia
{labaj,bielik}@fiit.stuba.sk

Abstract. Current educational systems offer many personalized features. But do they really help? Which personalization method is the best in given settings? Evaluating the personalization in an educational system is as important as designing the methods themselves. While many quantitative and qualitative methods have been explored previously, there are various rules and issues when performing experiments with participants – users. E.g. users should not interact with experimenters, but on the other hand, only post-session testing can be affected by maturation. We proposed a conversational evaluation approach in which we combine advantages of uncontrolled experiments with advantages of other methods. We describe a method for evaluation questions asked to the users in appropriate moments during their work in a web-based environment and its initial evaluation within ALEF adaptive learning framework.

1 Introduction and Related Work

When designing and operating adaptive and personalized systems in particular, evaluation is an essential part of the process. Without knowing the performance of the personalized system when employing various methods with various parameters, it is impossible to judge its effectiveness, pick successful methods, or make adjustments at all. Differing evaluation methods can be employed for one adaptive system, even on multiple levels, evaluating the adaptive system by parts [1], where we can often omit users, e.g. by using golden standards. Here we focus on those evaluation methods, which include users using the adaptive system being evaluated.

One can focus on if and how users interact with the experimenter in a user-centered evaluation and recognize several common approaches [1, 2]: *questionnaires* (series of questions displayed on paper or within a system), *interviews* (interviewer asks the user), *data log analysis* (user actions are recorded and analyzed without the participation of the user, *focus groups* (a discussion in a group of participants), and *think-aloud protocols* (the user describes their actions during the session).

Regardless of whether the users are interviewed, given pre-tests or questionnaires, or we only use logs of their activity, another categorization can made on how the experiment is performed. Three types of experiments are a common practice in adaptive systems such as recommender systems [3]: *offline* evaluation, *user studies* (*controlled experiments*), and *online* evaluation (*uncontrolled experiments*).

In *offline* experiments, previous user interaction with the system is recorded and used without the users. For example, we can record which learning objects were visited by users and how they were rated, predict user ratings using collaborative filtering and evaluate from the recorded data whether the user actually rated the learning objects as predicted. In a *user study*, a group of users in a controlled environment work with the system. The user feedback can be gathered using multiple methods – e.g. with think-aloud protocols during the session, while also using post-session questionnaires. In *online* experiments (where "online" does not indicate network environment, but rather live system being used), users work towards their goals in their own settings, for example a recommender system is deployed into live learning system.

There are rules which should be observed whenever possible, ranging from random assignment of participants, through instructing them in the same way, to maintaining uniform work environment [4]. These rules are aimed at leveling out the effect of nuisance variables and can be obeyed either by accordingly preparing the test room, written instructions, etc., or also by randomizing their influence, e.g. letting large volume of users work in their own environment at own times (naturally, doing the same for experiment and control groups).

One decision is that we either let the users go alone without any influences, and in fact, large number of users can participate this way, obtaining vast datasets for numerical evaluations, or we bring the users into laboratory, where we can have absolute overview of their actions, expressions, etc. and even talk to them (think-aloud methods). Whether we have the user at hand in laboratory, or perform the experiment remotely, when we seek out opinions from the users, or, in a learning system, want to assess their knowledge for evaluation using the measure of gained knowledge, we have several options. We can passively provide commenting or rating tools in the system and count on the users using them. We can interact with the user during the work (e.g. think-aloud), but this can alter the user's behavior, e.g. the user can approach problems differently when speaking out loud about them [4]. Or we can use post-testing, but we will maybe introduce maturation factors (users forget and also after the whole session, they can look differently on specific events).

We seek for a method of user-centered adaptive educational systems evaluation combining advantages of the above – capturing the user feedback right during users' workflow, but with minimal impact on learning process, i.e. without interrupting them significantly, and doing so in their natural settings. We proposed a conversational evaluation approach using *evaluation questions* being displayed at the appropriate times for collecting explicit user feedback. In this paper, we describe the evaluation questions approach and its use within our Adaptive LEarning Framework ALEF.

2 Conversational Evaluation

In our conversational evaluation approach, we generate *evaluation questions* and display them in appropriate moments during user's normal unsupervised work. Using evaluation questions is not unfamiliar to user feedback elicitation approaches, such as recommender rating elicitation [5], where an adaptive system asks the user for their ratings as needed, e.g., when the recommender does not have enough information to pick an item to be recommended for the given user.

Asking for the item rating instead of waiting for the user to provide one has several advantages: the users can be more motivated to even provide the rating in the first place, as they are told that they are helping improving their profile and helping the system with recommendation to others [5], the user can achieve higher satisfaction and perceive the system as more useful [6], and of course accuracy can be increased while decreasing the load on user, as ratings are elicited for such items whose the rating or re-rating would be useful to the system [7].

We follow similar line with conversational evaluation questions – instead of waiting for the user to provide the feedback after the session through the post-test/questionnaire, or on the other side, asking them to unnecessarily comment every thought, relevant evaluation questions are asked at appropriate moments. A user of an adaptive educational system can be sitting at home, studying for the next exam and at the same time helping evaluating/improving the system and associated personalization methods by answering sparsely displayed questions.

We propose a rule-based framework for generating conversations aimed at evaluation of user opinion, knowledge, etc. Evaluation questions are based on classic question types: yes/no, single choice, multiple choice, or free text. The text of questions is prepared by educational system developers (designer of particular personalization technique being tested) and stored as *questions templates*. When a question template is selected by the question engine and adapted to the user and situation (by processing the template scripts), it becomes a *question instance*, asked to a given user, within a given setup (e.g. learning course), and comes from a given asking rule.

The evaluation questions (their respective templates) are selected by question asking engine based on *triggers*. The triggers are composed of a rule-part with arbitrary conditions to be evaluated against the user model and of pre-assigned question template. For example, when a user scans through a list of items in a navigation tool (providing recommendations) and then proceeds to use the non-adaptive menu instead, a question template on why the tool was not used is triggered. The triggers have pre-assigned priorities and the most immediate question template takes over. As we are aiming at web-based systems, the triggers can be based on client-side user actions, as well as server side logs. Fig. 1 shows an overview of the involved entities.

The questions can be asked in a *synchronous* elicitation, which occurs during or just after an action, e.g. asking the user to rate a learning object after they finished reading this object, and in an *asynchronous* elicitation, which can, for example, occur when a user input is needed regardless of his actions.

The selected and instantiated question is displayed to the user in the foreground, darkening the rest of the system screen. In order not to obtain random answers (just for the window to go away), the user can chose not to answer this question, or to answer it later (if feasible due to the nature of the question). All these steps – from selecting the question template using triggers, to instantiating the question template, to asking and possibly re-asking the question, to receiving the answer(s), are logged.

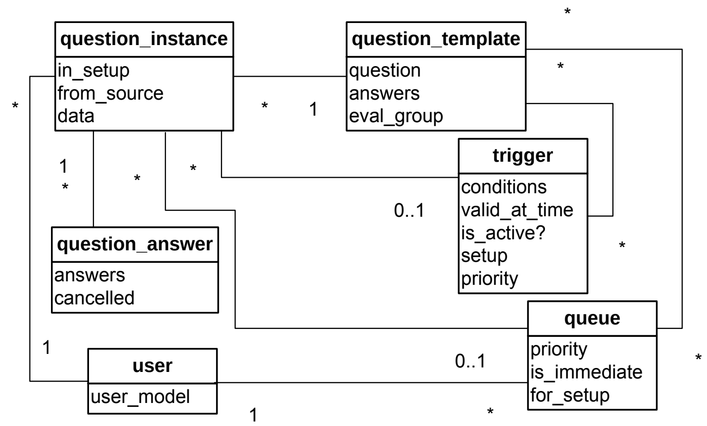

Fig. 1. Conceptual overview of the evaluation question model.

3 Evaluation Questions within ALEF Learning Framework

We realized the proposed evaluation questions approach as an evaluation sub-framework within the ALEF (Adaptive LEarning Framework) [8]. It is being used on the Faculty of Informatics and Information Technologies for five years in several courses: Principles of Software Engineering, Functional Programming and Logic Programming, and Procedural Programming. ALEF (Fig. 2) offers the users (students) learning objects of Explanation, Exercise, Question, and Question-Answer type, and also a variety of tools, ranging from personalized recommendations of learning objects, to automatic content enrichment using external sources. ALEF also provides domain and user models, as well as other features necessary for such personalized solutions. The educational system is used by students both during lectures in supervised settings in laboratories, as well as at home, unsupervised.

One example of a feature problematic to evaluate is the recommender system usage. When the user follows a recommendation to study the proposed learning object as next, it can be evaluated through measures such as time spent (immediate return is a negative indicator, staying for some time is a positive indicator that the user has liked the recommendation), through subsequent user rating of the item, or through adaptive knowledge testing via exercises and questions (user's knowledge has stayed the same or increased). However, when the user does not follow any recommendation (and this is a frequent case), it can have various meanings – maybe the user did not notice the recommendations, or he does not understand them, or the recommendations are inaccurate so the user ignores them, or the user just do not want to use them at all for own reasons often related to personal goals and motivational elements for using the educational system. Do we need to make the recommendations visually more prominent in the system? Or explain them in a better way? Employ different recommendation method? Questions like these can be answered by prompting the user non-invasively during his focus changes when using other tools in the system to navigate (e.g. ask: Why did he choose the menu over the recommended items?).

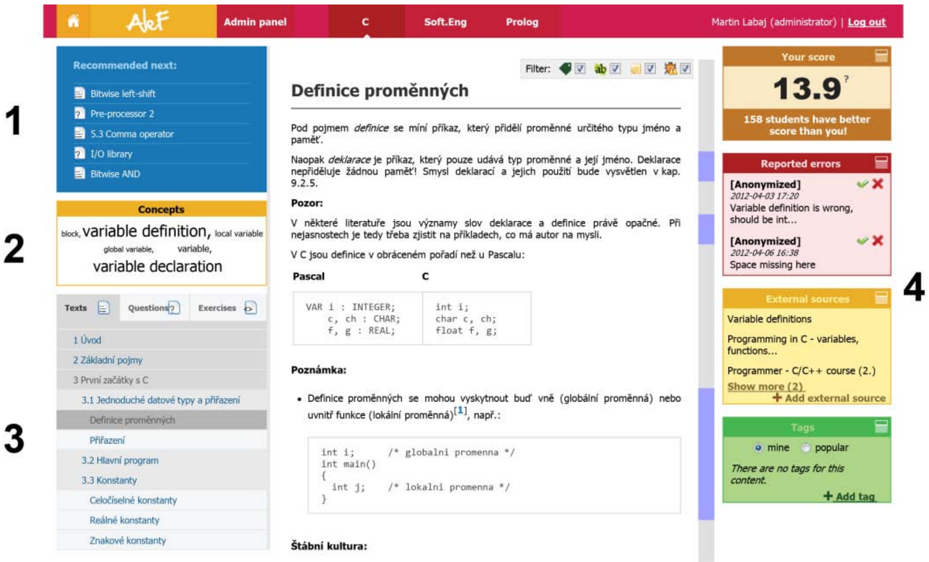

Fig. 2. Screenshot of ALEF educational system, showing a learning object ("Variable definition", in Slovak) in the middle. The left side contains tools for navigation between learning objects: personalized recommendations (1), tag recommendation (2), menu (3). The right side contains tools within the learning object (4): reported errors, external sources, tags.

4 Evaluation and Conclusions

Our rule-based evaluation question framework was used in several evaluations, most recent on summarization of explanation parts of learning objects [9]. Here we present a case of synchronous questions based on user attention. 34 students took part in two week uncontrolled experiment. User attention was tracked using mouse interaction and commodity gaze tracking via webcams and based on attention, application (tools) and document (learning object) fragments were assessed and recommended.

We hypothesized in this experiment that when asked at the appropriate moments, the users are more willing to provide opinions. The questions were displayed by the question engine in two situations. When the user focused on a fragment for a period of time and then shifted focus away, a question about the fragment was instantiated. The same question templates were also triggered randomly, unrelated to user attention, i.e. even during focused work, and unrelated to their current target fragment. In each question, the users had access to afore mentioned options of postponing the question or declining to answer. When asking questions related to user's previous fragment and in the moments of shifting the focus, only 7 % (using gaze and mouse) and 12 % (mouse only) of the instantiated questions were cancelled. Contrast to this, 33 % of randomly displayed questions was dismissed.

Our experiment suggests that when the evaluation questions are asked at the appropriate time and right when the user is working with the part in question (or just finished working with it), we can accomplish higher cooperation from the user providing more feedback than when we would ask them randomly. This is not the only advantage. Since the users provide their opinions during their work, right when they interact with a given object, such as recommendation, they provide higher quality feedback than when commenting/rating after the entire session, and yet, they are not as interrupted as in supervised think-aloud evaluation.

Our approach does not aim to entirely replace the physical presence of the user and the possibility of observing directly what are they doing and asking additional, unprepared questions. It is rather a supplement, which gives different views on adaptive mechanisms and collects such views from users participating even in an unsupervised online experiment. Although the questions are constructed when asked and adapted to the user, some pre-thought is needed to create the templates in advance.

Currently we are interested in combining the conversational evaluation with the elicitation of ratings, especially based on user attention and also in continuous adaptive testing of user knowledge, using these mechanisms with questions created by the teacher or sourced from the learning content.

Acknowledgements: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/0971/11 and by the Slovak Research and Development Agency under the contract No. APVV-0233-10.

References

1. Paramythis, A., Weibelzahl, S., Masthoff, J.: Layered evaluation of interactive adaptive systems: framework and formativemethods. User Modeling and User-Adapted Interaction. 20, 383–453 (2010).
2. Van Velsen, L., Van Der Geest, T., Klaassen, R., Steehouder, M.: User-centered evaluation of adaptive and adaptable systems: a literature review. The Knowledge Engineering Review. 23, 261–281 (2008).
3. Shani, G., Gunawardana, A.: Evaluating Recommendation Systems. In: Ricci, F., Rokach, L., Shapira, B., and Kantor, P.B. (eds.) Recommender Systems Handbook. pp. 257–297. Springer US, Boston, MA (2011).
4. Chin, D.: Empirical evaluation of user models and user-adapted systems. User modeling and user-adapted interaction. 181–194 (2001).
5. Carenini, G., Smith, J., Poole, D.: Towards more Conversational and Collaborative Recommender Systems. Proceedings of the 8th international conference on Intelligent user interfaces - IUI ’03. pp. 12–18. ACM Press, New York, New York, USA (2003).
6. Knijnenburg, B.P., Willemsen, M.C.: Understanding the effect of adaptive preference elicitation methods on user satisfaction of a recommender system. Proc. of the third ACM conf. on Recomm. sys. - RecSys ’09. p. 381. ACM Press, New York, USA (2009).
7. Amatriain, X., Pujol, J.M., Tintarev, N., Oliver, N.: Rate it Again: Increasing Recommendation Accuracy by User re-Rating. Proc. of the third ACM conf. on Recommender. systems - RecSys ’09. pp. 173–180. ACM Press, New York, USA (2009).
8. Šimko, M., Barla, M., Bieliková, M.: ALEF: A framework for Adaptive Web-Based learning 2.0. Key Competencies in the Knowledge Society, WCC '10. pp. 367–378 (2010).
9. Móro, R., Bieliková, M.: Personalized Text Summarization Based on Important Terms Identification. 23rd Int. Workshop on Database and Expert Systems Applications. pp. 131–135. IEEE (2012).

Exploiting Human Signals in Learning Environment as an Alternative to Evaluate Education Performance

Fernando Koch, Marcia Ito, Annibal B. M. da Silva, Sergio Borger, and Julio Nogima

IBM Research - Brazil
email: {ferkoch, marcia.ito, abiondi, sborger, jnogima}@br.ibm.com

Abstract. There is a demand for new ways to understand the relation between student and group behaviour, and their impact on education performance. For that, we are researching, developing, and testing technologies to instrument classrooms, collect human signal data, and derive meaning that leads us to understand their relation with the education performance. We call this setup as "the Smarter Classroom". It integrates (i) applications running on tablet computing devices that play digital education content and collect students' gestures whilst manipulating the materials, (ii) environmental sensors such as video cameras and microphones, and (iii) innovative Analytics models that can make sense of these signals. In this work, we describe our development, present a practical experiment, and discuss the field applicability of this technology.

1 Introduction

The role of the modern education system is to provide students with skills and knowledge to prepare them to pursue advanced degrees and employment to be able to succeed in a globally competitive world [5]. This means that institutions must tailor learning experiences to their students towards the ideal of massification with personalisation of the education process. For that, there is a demand for new ways to understand the relation between student and group behaviour, and their impact on education performance.

We are developing learning environments that collect and store the *human signals* [10] generated during the learning process. We call this development as "the Smarter Classroom". It provides comprehensive and affordable instrumentation of classrooms along with innovative Analytic models that can make sense of this data. For example, we analyse signals like the time spend on a page, clicks, zooming gestures, taps, ambient sound, disturbances in the classroom, and others. Based on this information we can deduce individuals' behaviours like interest, attention, focus thought, and others [9], as well as insights on group behaviour.

The solution integrates (i) applications running on tablet computing devices that play digital education content and collect students' gestures whilst manipulating these materials, (ii) environmental sensors such as video cameras and microphones, and (iii) Analytics models to make sense of the data being collects. For the latter, we are exploiting the concepts of *Social Analytics* [1] and *Learning Analytics* [4] aiming to create the intelligence to:

- Classify individual and group behaviour based on human signals in learning environments.

– Correlate social behaviour to education performance.
– Recommend actions to improve the education performance, as for example adjustments in the learning environment, modifications in he content, distribution of students based on social roles, and others.

The paper is structured as follows. Section 2 describes the motivation and related work. Section 3 presents the prototype implementation and practical experiments. The paper concludes with Section 4 with an analysis of the results and a discussion about the field applicability of this technology.

2 Motivation and Related Work

We aim at tools to integrate the pillars of the education environment, *i.e.* teachers, students, the classroom, and planning. Our proposal is to create new methods to track and evaluate the students' performance taking in consideration how they interact with the education material, and with other students.

In the field of *Ambient Intelligence*, the work by [8] introduces an integrated architecture for pervasive computing environments in *Project ClassMATE*. The work in [11] proposes the use of sensors and speech recognition integrated to an analysis model in *project iClass*. The report in [2] discusses the opportunities and consequences of applying these techniques in the classroom environment.

Related to *Learning Analytics*, the report in [4] presents diverse approaches for the measurement, collection, analysis and reporting of data about learners and their contexts. The work in [12] provides a broad view of the use of Analytics in education environments. The work in [3] introduces Social Learning Analytics by combining learning analytics and social networks.

Moreover, we are motived by the work in [9], where human signals are collected and analysed to read people, allowing to classify individual and group behaviour, social roles, patterns in group interactions, and the development of social networks, and others.

The related work identified in the prior art provide the basis for the study being conduced in this project. We seek an integrated solution that exploits the concepts of data collection and environment iteration in *Ambient Intelligence* and the methods to extract deep insights provided by *Learning Analytics*. However, we want to use human signals as the reference data – instead of simply using exams' marks or surveys like usual analytic models in the latter. Hence, we identified an opportunity to contribute with a combined model as outlined below.

3 Prototype and Experiment

Figure 1 depicts the solution overview of the Smarter Classroom. It contains (1) *Front-end solutions* to instrument classrooms environment, *e.g.* with video cameras, voice capturing, ambient sound capturing, and applications running on tablet computing devices that play digital education content. For instance, we prepared a scenario where the teacher is equipped with a headset and a tablet computing device with a special control application. The teacher's voice is streamed to an Automated Speech Recognition

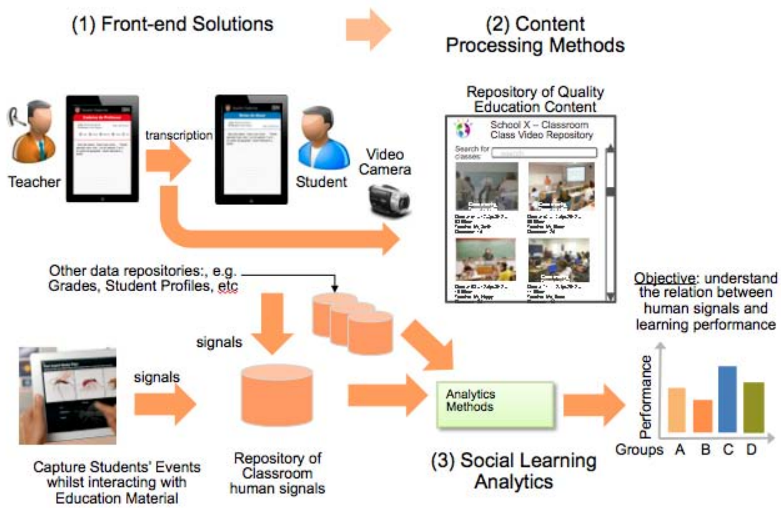

Fig. 1. Solution Overview: Smarter Classroom

service, generating the transcription. This is streamed to the students' tablet computing devices, creating a on-line annotation system. We also instrumented the classroom with a camera so to record the class. The content player was developed as a version of the Cool Reader [7], which is an open source e-book reader for Android. It is capable to handle standard formats like EPUB and FictionBook. The code has been instrumented to capture the signals, and store the data in log files. We represent a signal s_i as the tuple $< ts, tp, pr >$ where ts is the timestamp, tp is the type (*e.g.* page turn, zoom in, zoom out, link clicked, others), and pr are description parameters. At the end of the class, the applications upload the log files to a server where they are stored and indexed.

The environment also provides (2) *Content processing methods* to compile the captured data, making it available to other systems, students and administrators. In the sideline, we are exploiting this module to integrate with Content Management Systems in order to create dynamic web sites and repositories of quality education material.

Finally, the (3) *Social Learning Analytics methods* implements the models to derive meaning from the collected data. It works by a combination of calculation models in form of mathematical and statistical functions that process the human signals captured by the (1) *Front-end Solutions*.

For instance, let us say that: $M = \{m_1, \ldots, m_n, t_1, \ldots, t_m\}$ is the education material composed of the set of elements m_i (*e.g.* (text, figures, links, etc) and multi-choice test t_j, and the $S_{\{c,M\}} = \{s_1, \ldots, s_n\}$ contains the signals captured from a student c using M. The classroom $C = \{c_1, \ldots, c_n\}$ is a set of students. Then, we developed calculation models as for instance:

– *Calculate level of activity while resolving a task*: given a task to read elements and respond to tests $I \subseteq M$; there is a function $levAct(S_{\{c,M\}})$ that calculates the level of activity ac_c whilst resolving the task; for instance, a calculation of time between groups of events; there is a function $avgAct(C) \rightarrow \alpha$ that calculates the average level of activity of the students in C. The function $act(c, I)$ classifies level of activity as: *slow activity* if $ac_c \leq \alpha * (1 - T_{ac})$, *normal activity* if $\alpha * (1 + T_{ac}) > ac_c > \alpha * (1 - T_{ac})$, and *high activity* if $ac_c \geq \alpha * (1 + T_{ac})$, where T_{ac} is a threshold (*e.g.* $T_{ac} = 0.2$ in our experiments).

– *Calculate level of attention while resolving a task*: given a task to read elements and respond to tests $I \subseteq M$; there is a function $levAtt(S_{\{c,M\}})$ that calculates the level of attention at_c whilst resolving the task; for instance, it takes in consideration the time between actions, time switching in and out the application (*i.e.* distractions by other applications), and others; there is a function $avgAtt(C) \rightarrow \beta$ that calculates the average level of attention of the students in C. The function $att(c, I)$ classifies level of activity as: *inattentive* if $at_c \leq \beta * (1 - T_{at})$, *attentive* if $\beta * (1 + T_{at}) > at_c > \beta * (1 - T_{at})$, and *highly attentive* if $at_c \geq \beta * (1 + T_{at})$, where T_{at} is a threshold (*e.g.* $T_{at} = 0.5$ in our experiments).
– *Calculate performance resolving a task*: given a task to read elements and respond to tests $I \subseteq M$; there is a set $E(M, I) = \{e_1, ..., e_n\}$ of optimal sequence of events to execute the instruction; there is a function $distOpt(S_{\{c,M\}}, E(I))$ that calculates the inverse of the distance pf_c between the sequence executed by the student and what would be the optimal sequence; there is a function $avgDist(C) \rightarrow \delta$ that calculates the average performance of the students in C. The function $perf(c, I)$ classifies performance as: *low performance* if $pf_c \leq \delta * (1 - T_{pf})$, *normal performance* if $\delta * (1 + T_{pf}) > pf_c > \delta * (1 - T_{pf})$, and *high performance* if $pf_c \geq \delta * (1 + T_{pf})$, where T_{pf} is a threshold (*e.g.* $T_{pf} = 0.2$ in our experiments).

Finally, there is a method to *Calculate performance resolving tests* based on the number of right answers provided to the tests $\{t1, \ldots, t_n\} \subset I$. We can then implement experiments to collect data and apply these methods in order to classify individual and group behaviour in learning environment as demonstrated below.

3.1 Experiment

In this experiment we focused on the detection of Attention Deficit Hyperactivity Disorder (ADHD) and analyse their impact in education performance. Our hypothesise is that depending on the students' behaviour it is possible to classify their profiles as ADHD inattentive, ADHD hyperactive, or normal behaviour and then compare the results from observations based on surveys conduced with these students.

We implemented a subset of the Smarter Classroom – *i.e.* tablet computers with the player application and digital education material – in a controlled environment containing students with diverse profiles[1]. The teacher delivers the class explaining in detail the whole digital education material M. Next, the teacher requests the students to execute a set of tasks to find the elements of $I \subset M$. The students execute these activities, generating logs $S_{s,M}$. Table 1 presents example results.

[1] ADHD detection: as there is no final diagnosis for ADHD level, the students have been individually evaluated based on their self-classification and behaviour.

Table 1. Example of Test Results

	Low Activity	Normal Activity	High Activity
Inattentive	Task Low / Exam Low	Task Medium / Exam Low	Task Low / Exam Low
Attentive		Task Medium / Exam Medium	Task High / Exam Medium
Highly Attentive		Task High / Exam High	Task High / Exam High

From the results, we notice that students classified as *inattentive* whilst utilising the education material attain lower performance for both task execution and exams. We concluded that the students with *low activity* in this group present the characteristics of ADHD inattentive, whilst the ones with *high activity* tend towards ADHD hyperactive – however, we grant that this observation is not conclusive and may not be always the case. During the survey, the students with known ADHD inattentive condition reported difficulty to: pay attention to the class, understand what is being discussed in a given moment, and keep attention whilst the tablet computing offers other distractions (*i.e.* applications other than the content player). On the other hand, the students with known ADHD hyperactive condition reported that they need to feel in control of the tablet computing and player application, so they spent considerable amount of time playing with the configurations. Some reported problems with the application (most likely due to misconfiguration), which let them feel impatient and disappointed with the technology.

Conversely, students classified as *attentive* and *highly attentive* attain best performance in both metrics. We cannot conclude that high activity in manipulating the education content necessarily reflects ADHD conditions for these groups. During the survey, the normal students (*i.e.* the ones whose ADHD condition is not detected) reported that: "it was easy to use the player application and the interface is friendly". Some of the *highly attentive* users complained that other students were taking too long to complete the tasks, delaying their performance in class.

This experiment demonstrate the feasibility and potential of the technology. It is missing now more Analytic modes able to computer different performance indicators and apply the technology in diverse and larger environment to validate the results.

4 Conclusions

We presented our research in creating a interface to capture human signals in learning environment, integrated to innovative analytic models to extract meaning from this data. This development leads to alternative methods to classify and understand the impact of individual and social behaviour in the learning environments. We acknowledge the legislative, ethical, and organizational issues related to the field implementation of this proposal. However, so far we are working on proving the concept and applicability of the solutions. In further stages we will discuss the practices for field implementation.

We implement a proof-of-concept experiment to detect variations of attention deficit hyperactivity disorder (ADHD) based on level of attentiveness, activity and task performance. We could successfully detect the human signals involved in this situation and related to performance and activity whilst resolving education tasks. This experiment demonstrates the feasibility and potential of applying this technology in the field.

This development advances the state-of-the-art by introducing a method to analyse education performance based on patterns in human signals. We are building upon the solutions and case scenarios in the *IBM Smarter Education* program [6], which envisages the use of analytics to understand the learning environment. We aim to contribute to this program with a layer of understanding about individual and group behaviour and its impact on education performance.

Future work will provide extended analytic methods, implement larger test scenarios, and create recommendation modules and visualisations to facilitate decision making. In the long term, we aim to integrate these modules in a composed solution.

Acknowledgement This work has been supported and partially funded by FINEP / MCTI, under subcontract no. 03.11.0371.00.

References

1. C. Aggarwal. *Social Network Data Analytics*. Springer Publishing Company, Incorporated, 1st edition, 2011.
2. J. C. Augusto. Ambient intelligence: Opportunities and Consequences of its Use in Smart Classrooms. *Italics*, 8(2):53–63, 2009.
3. S. Buckingham Shum and R. Ferguson. Social Learning Analytics. *Educational Technology & Society*, 15(3):3–26, 2012.
4. R. Ferguson. The State of Learning Analytics in 2012: A Review and Tuture Challenges. Technical Report KMI-2012-01, Knowledge Media Institute, 2012.
5. A. Green. *Education, Globalization and the Nation State*. ERIC, 1997.
6. IBM Corp. IBM Smarter Education. http://www.ibm.com/smarterplanet/education, last checked May-2013.
7. V. Lopatin. Cool Reader 3. http://coolreader.org/e-index.htm, last checked May-2013.
8. G. Margetis, A. Leonidis, M. Antona, and C. Stephanidis. Towards Ambient Intelligence in the Classroom. In *Proceedings of the 6th international conference on Universal access in human-computer interaction: applications and services - Volume Part IV*, UAHCI'11, pages 577–586, Berlin, Heidelberg, 2011. Springer-Verlag.
9. A. Pentland. *Honest Signals: How They Shape Our World*. The MIT Press, 2008.
10. A. Pentland. To Signal is Human. *American Scientist*, 98(3):204–211, 2010.
11. R. A. Ramadan, H. Hagras, M. Nawito, A. Faham, and B. Eldesouky. The Intelligent Classroom: Towards an Educational Ambient Intelligence Testbed. In *Intelligent Environments (IE), 2010 Sixth International Conference on*, pages 344–349, 2010.
12. G. Siemens and P. Long. Penetrating the Fog: Analytics in Learning and Education. *Educause Review*, 46(5):30–32, 2011.

Open Issues in Educational Affective Recommendations for Distance Learning Scenarios

Ángeles Manjarrés-Riesco, Olga C. Santos, Jesus G. Boticario, Mar Saneiro

aDeNu Research Group. Artificial Intelligence Dept. Computer Science School. UNED

C/Juan del Rosal, 16. Madrid 28040. Spain

{amanja,ocsantos,jgb,marsanerio}@dia.uned.es

Abstract. Despite psychological research showing that there is a strong relationship between learners' affective state and the learning process, affection is often neglected by distance learning (DL) educators. In this paper we discuss some issues which arise when eliciting personalized affective recommendations for DL scenarios. These issues were identified in the course of applying the TORMES user centered engineering approach to involve relevant stakeholders (i.e. educators) in an affective recommendation elicitation process.

Keywords: Educational recommender systems, affective computing, distance learning, educational scenarios.

1 Introduction

Recommender systems should help and support both learners and educators in educational web based scenarios [1]. Very often, this support has been given in terms of content to be read by the learners to reduce the information overload (see relevant compilations of educational recommenders in [2-6]), while educational scenarios might offer richer recommendation opportunities that involve the usage of learning services [7]. Moreover, psychological research shows that there is a strong relationship between the learners' affective state and the learning process [8-9]. However, to date there have been only a few recommender systems in educational scenarios that have considered affective issues. They have been used to 1) recommend courses according to the inferred emotional information about the user [10], 2) customize delivered learning materials depending on the learner emotional state and learning context [11] and 3) provide the list of most suitable resources given the learner affective state, provided that the learner fills in i) her current affective state (flow, frustrated, etc.) and ii) her learning objectives [12].

In the past, we have been researching on eliciting educational oriented recommendations to identify recommendation opportunities that go beyond reducing the information overload in learning management systems. Thus, in order to support educators in the elicitation process, we proposed the TORMES methodology [14]. TORMES adapts the ISO standard 9241-210 to help educators identifying when, who, what, how, where and why educational support needs to be provided to each particular learner in a given educational scenario, as well as on which features characterize the recommendations. When we came to take into account affective issues in the learning process within the MAMIPEC project [29], TORMES methodology was extended to explicitly support educators in eliciting recommendation opportunities that involved emotional feedback [25]. An initial application of TORMES was done with 3 educators particularly concerned with teaching strategies that incorporate the affective dimensions. Moreover, additional 12 educators, who have not taken part in the elicitation process, were asked to evaluate 12 of the 47 affective scenarios elicited and their corresponding potential recommendations. Preliminary outcomes of this application have been reported elsewhere [27, 28]. In general terms, educators who evaluated the recommendations elicited found them as valuable affective pedagogical interventions. However, some open issues were identified. In some cases, educators pointed out that applying recommendations into real practice was beyond their capabilities. In particular, they reported difficulties in detecting the need of affective support in real learning scenarios, which in our view, shows that DL educators might not intervene in certain valuable affective ways due to the lack of resources and training related to dealing with the student affective state and applying appropriate intervention strategies.

In this context, the goal of this paper is not to report on experiment findings, but to reflect on the work done so far and identify open issues relevant to affective recommender system research to be shared and discussed during the workshop with the rest of the participants. In this way, next we present the open issues identified and after that, we comment on how, in our view and experience, they can be addressed.

2 Identified open issues concerning affective educational support in DL

From our past experience, review of the pertinent literature, and outcomes from the application of TORMES, we have identified 5 open issues that mainly involve a lack of resources and training concerning affective teaching on the part of DL educators.

2.1 Scarce reported experiences on affective support in DL scenarios

Although general models of affective support in e-learning have been proposed (as in the case of [13]) and some positive studies have been reported (such as [15]) on the commitment of DL institutions with the principles of an "affective teaching", to date, affection is often neglected by DL educators. Educators have mainly focused on the cognitive domain of learning [30] and they have ignored its affective domain. As consequence educators are poorly trained in affective teaching strategies. In fact, there is no literature on the term "affective teaching" itself, but only on affective learning, and the significant literature about academic emotions [16-18] focus on face to face learning experiences usually concerning students of secondary and higher education. In any case, it is acknowledge the benefits of providing a positive emotional climate

[36] where learners are happy and feet well supported in their learning [37, 38], since a climate built on mutual trust could encourage learners to take on new learning challenges as they would not be afraid of making mistakes [38].

However, distinctive and unique affective experience issues intricately linked to the computer interaction experience (supported by e-learning platforms) concern DL students. In our view, their singularity and the distance context itself should give rise to particular educational scenarios and affective responses that require particular approaches of student affective support. Thus, we consider of importance to analyze the diversity of scenarios with affective relevance that may arise in DL contexts.

2.2 Affective needs in DL

There is abundant literature characterizing distance learners [21-22]. Learning provide new stimulus to these students, guided by intrinsic motivations, and determine the mobilization of intense emotions. But it is essential to consider that learning for adult students entails different characteristics to those belonging to other population groups. Adult students feel less fitted, and tiredness and lack of time are consequences of their socio-occupational status. Learning requires great personal sacrifices that do not prevent very long study times. Demotivation is the main cause of dropout. The quality of electronic communications is not enough to satisfy socialization needs, and causes conflicts with affective implications [23-24]. Taking into account these factors educators should consider how affective issues could impact on the learning process [31], helping learners to recognize and manage their own emotions, by increasing motivation, facing critics, etc. Affective learning is subject of consideration particularly in the case of students with disabilities, who tend to choose the distance modality and thus difficulties caused by their own characteristics should be considered when they are communicating or understanding emotions showed by others. Sometimes disabilities involve deficits in the different stages of affective processing (sensing, expressing, or interpreting affect-relevant signals). Consequently, people with these kinds of disabilities can be considered emotionally handicapped [39]. Besides, adult students deserve a treatment different from young face-to-face students, demanding more participative teaching approaches, subtle and suggestive support, respect and appreciation of their experience, further reinforcement and motivation, friendliness and closeness.

2.3 Difficulties of affective communication in virtual learning communities

In DL, given the lack of straightforward information on student affective states, this is inferred from various sources, such as forum and email messages, as well as occasional telephone calls that express emotions more or less directly. Frequency of learners' communications and interactions in virtual courses may also indicate hidden emotional states. There is no doubt that it is difficult to assess with certainty the emotions involved, their intensity, their permanency, etc. only from these information sources, providing they facilitate emotional dishonest information about their feelings and emotions because they feel more vulnerable and incapable. Moreover, the educator must express affection to her students to dissolve the students' natural tendency to resist being told what to do, so the advice can penetrate more deeply and effectively [34]. Humor can be use as a tool to make the topic or subject seem more relevant to learners' own experiences [36], as it can engender trust and mutual respect and made every effort to be flexible in order to provide a learning environment that encouraged pupil participation [37].

Another issue here, is the difficulty of affective communication influences not only learning itself, but social relationships in virtual communities [18-20], among others. Virtual learning communities often are a meeting place between students and educators. Accordingly, social relationships are a key aspect of learning since if the learner has not adequate social competences, she will not be able to acquire and share information and knowledge with others, receive feedback about her beliefs, work, etc., impeding her to modify or improve them.

2.4 Reduced scope of the affective support provided in current approaches

In the last century different authors have examined the domains of learning and they identified levels of learning in affective domain. The affective domain covers motivation, emotions, values, attitudes and behaviors [32]. Thus, affective support should take into consideration student psychological factors such as attitudes, beliefs, motivation and thoughts. However, current affective research focuses almost exclusively on increasing learners' motivation. Thus, no emphasis is done on providing learners with emotional regulation strategies for the benefit of the learning process. These strategies include activities and resources to improve the ability to listen, demonstrate attitudes, revise judgments and change an inadequate behavior and could be provided through educational oriented recommendations focused on recommending specific actions to be carried out by the learners.

2.5 Lack of resources for educators to provide affective support

In general educators, and more specifically distance educators, face difficulties when teaching affective outcomes, they consider that these issues are private and far too long term to be integrated into any learning program [33]. As consequence, distance education teachers are not usually aware of the impact of affection in learning, and are not used to provide affective support to their learners, despite the fact that underlying any instruction there is always an implicit affective support strategy. Therefore, a methodology is needed to help educators elicit recommendation opportunities in their teaching scenarios. However, we could not find in the literature of educational recommender systems [2-6] methodological approaches to support the recommendations elicitation process except for the TORMES methodology that we have applied.

3 Discussion on ways to address the open issues

Educational recommender systems can model the affective issues involved during the learning process, considering that this modeling has to be managed and integrated with the rest of existing e-learning services. Given the open issues in affective learning theories, the heuristic knowledge that is applied in everyday instruction practice in learning institutions might be of great importance. As for the current literature on this topic, large parts of this knowledge have not yet been collected.

Moreover, in our view, there is a lack of methodological approaches to support the recommendation elicitation process in user modeling and personalized educational scenarios. This need is even more critical in distance teaching, where affective support would be very valuable but has been usually neglected. In view of the above, we propose the involvement of educators in order to carry out an exhaustive and methodical compilation of heuristics concerning affective learning in DL contexts, as already suggested in the literature (e.g., see [8]), by applying TORMES.

As shown, the TORMES methodology can be of help to support eliciting recommendation opportunities in which affective issues can be addressed to support the learning process from educators [25]. In past experiments [26]), which did not focus on affective issues, we already found a statistically significant positive impact on indicators dealing with engagement in the course, learning effectiveness and efficiency, and knowledge acquisition when educational recommendations are delivered to learners in the learning management systems.

From the evaluation activity we carried out [28], it appeared that there is little awareness and little training regarding affective educational dimension but a latent sensibility to the issue. It will be therefore advisable to extend and reformulate the elicited recommendations in the light of an affective teaching model that incorporates the theory and experiences of face-to-face courses translated to a DL context. Moreover, we belief necessary deepen the rich emotional universe of DL students by also engaging them in an affective scenario elicitation process.

Finally, the limitations on affective communication in DL scenarios are more difficult to address. We are currently working hard in the automatic detection of emotions on the basis of physiological parameters, but we are aware of the risks of misinterpretations inherent in context-aware system approaches, mainly when they involve such complex factors as emotions.

In summary, integrating affective recommender systems in e-learning platforms could contribute to raising awareness and training for an affective teaching. Thus, these systems could provide undoubtedly added value to e-learning platforms.

Acknowledgments

Authors would like to thank the Spanish Ministry of Economy and Competence for funding the MAMIPEC project (TIN2011-29221-C03-01).

References

1. Zaïane, O.R. Building a Recommender Agent for e-Learning Systems. Proceedings of International. Conference on Computers in Education, p. 55-59, 2002.
2. Manouselis, N., Drachsler, H., Verbert, K., Santos, O.C. Proceedings of the 1st Workshop on Recommender Systems for Technology Enhanced Learning. Elsevier's Procedia Computer Science, vol 1(2), 2010.
3. Manouselis, N., Drachsler, H., Verbert, K., Santos, O.C. Proceedings of the 2nd Workshop on Recommender Systems for Technology Enhanced Learning. CEUR-WS proceedings, vol. 896, 2012.
4. Manouselis, N., Drachsler, H., Verbert, K., Duval, E. (2013) Recommender Systems for Learning. Springer Briefs in Electrical and Computer Engineering.
5. Santos, O.C., Boticario, J.G. (2012) (Eds.) Educational Recommender Systems and Techniques: Practices and Challenges. IGI Publisher.
6. Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., Duval, E. Context-Aware Recommender Systems for Learning: A Survey and Future Challenges, IEEE Transactions on Learning Technologies, vol. 5, no. 4, pp. 318-335, 2012.
7. Berthold, M., Pachtchenko, S., Kiefel, A., Nussbaumer, A., Albert, D. Identifying Requirements for a Psycho-Pedagogical Mash-up Design for Personalising the Learning Environment. In: Diana Perez-Marin, Milos Kravcik, Olga C. Santos (Eds.) Proceedings of the International Workshop on Personalization Approaches in Learning Environments, held in conjunction with the 19th User Modeling, Adaptation and Personalization conference (UMAP 2011), 36-40, 2011.
8. Porayska-Pomsta, K., Mavrikis, M., Pain, H. Diagnosing and acting on student affect: the tutor's perspective. User Model. User-Adapt. Interact. 18(1-2): 125-173, 2008.
9. Blanchard, E.G., Volfson, B., Hong, Y.J. Lajoie, S.P. Affective Artificial Intelligence in Education: From Detection to Adaptation. In Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling (AIED 2009), 81-88, 2009.
10. Gonzalez, G., de la Rosa, J.L., Montaner, M.., Delfin, S. Embedding Emotional Context in Recommender Systems. In Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop (ICDEW '07). p. 845-852, 2007.
11. Shen, L., Wang, M., Shen, R. Affective e-Learning: Using "Emotional" Data to Improve Learning in Pervasive Learning Environment. Educational Technology & Society (ETS) 12(2), 176-189, 2009.
12. Leony, D., Pardo, A., ParadaGélvez, H.A., Delgado Kloos, C. A widget to recommend learning resources based on the learner affective state. 3rd International Workshop on Motivational and Affective Aspects of Technology Enhanced Learning (MATEL), 2012
13. Al Masum, M., Wang, H., Mitsuru, I., Al Mamun, M.. Modeling an Affectionate Virtual Teacher for e-Learning underpinning 3-Dimensional Emotion Model. International Conference on Computer and Information Technology. 2005.
14. Santos, O.C., Boticario, J.G. TORMES methodology to elicit educational oriented recommendations. Lect. Notes Artif. Intell. Vol. 6738, p. 541–543, 2011.
15. Fonseca A., Salete Leone Ferreira, B. , Fernandes Gonçalves, S., Toledo Coutinho, R.E.. Affectivity in the Process of Distance Technical Education. Volta Redonda – RJ – Brazil - May 2011.
16. Kim, C.M., Hodges, C.B. Effects of an Emotion Control Treatment on Academic Emotions, Motivation and Achievement in an Online Mathematics Course. Instructional Science: An International Journal of the Learning Sciences, vol. 40 (1), 173-192, 2012.

17. Pekrun, R., Frenzel, A.C., Götz, T., Perry, R.P. The control-value theory of achievement emotions: an integrative approach to emotions in education. In: Emotion in education. Schutz, P.A. and Pekrun, R. (eds.). Academic Press, 13-36, 2007.
18. Pekrun, R., Elliot, A.J., Maier, M.A. Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. Journal of Educational Psychology, vol. 101(1), 115-135, 2009.
19. Wolf, A. Emotional expression online: Gender differences in emoticon use. CyberPsychology amd Behavior vol. 3(5), p. 827-833, 2000.
20. Ahn, J., Borowiec, A., Buckley, K., Cai, D., Chmiel, A., Czaplicka, A., Dąbrowski, G., Garas, A., Garcia, D., Gobron, S., Hillmann, R., Hołyst,J., Kappas, A., Küster, D., Mitrovic, M., Paltoglou, G., Pirker,H., Rank, S., Schweitzer, F., Sienkiewicz, J., Skowron, M., Sobkowicz, P., Thalmann, D., Thelwall, M., Theunis, M., Trier, M., Tsankova, E., Weronski, P., CYBEREMOTIONS – Collective Emotions in Cyberspace, Procedia Computer Science, vol. 7, p. 221-222, 2011.
21. Holmberg, B. Distance Education in Essence – an overview of theory and practice in the early twenty-first century. 2nd edition. Oldenburg: Bibliotheks- und Informationssystem der Carl von Ossietzky Universität Oldenburg, 2003.
22. Tennant, M., Pogson, P. Learning and change in adult years. A developmental perspective. San Francisco: Jossey-Bass, 1995.
23. Kreijns, K., Krischner, P. A., Wim, J. Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. In Computers in Human Behaviour, 19, 335-353, 2003.
24. O'Reagan, K. Emotion and E-learning. In Journal of Asynchronous learning networks 7(3), 2003.
25. Santos, O.C., Boticario, J.G. Affective Issues in Semantic Educational Recommender Systems. In Proceedings of the 2nd Workshop on Recommender Systems for Technology Enhanced Learning. CEUR-WS proceedings, vol. 896, 71-82, 2012.
26. Pascual-Nieto, I., Santos, O.C., Perez-Marin, D., Boticario, J.G. Extending Computer Assisted Assessment systems with Natural Language Processing, User Modeling, and Recommendations based on Human Computer Interaction and Data Mining. In Proceedings of the International Joint Conference on Artificial Intelligence, Barcelona, Spain, 19–22 July 2011; 2519-2524, 2011.
27. Manjarrés-Riesco, A., Santos, O.C., Boticario, J.G. Eliciting Affective Recommendations to Support Distance Learning Students. In proceedings of the 21st Conference on User Modeling, Adaptation and Personalization (UMAP 2013), 2013, in press.
28. Santos, O.C., Boticario, J.G., Manjarrés-Riesco, A. An approach for an Affective Educational Recommendation Model. Recommender Systems for Technology Enhanced Learning: Research Trends & Applications. Manouselis, N., Drachsler, H., Verbert, K., Santos, O.C. (eds.). Springer, 2013 (in press).
29. Santos, O.C., Boticario, J.G., Arevalillo-Herraez, M., Saneiro, M., Cabestrero, R., del Campo, E., Manjarrés_Riesco, A., Moreno-Clari, P., Quiros, P., and Salmeron-Majadas, S. MAMIPEC - Affective Modeling in Inclusive Personalized Educational Scenarios. In Bulletin of the Technical Committee on Learning Technology. Special Issue Articles: Technology-Augmented Physical Educational Spaces, vol 14 (4), October, 35-38, 2012.
30. Shepard,K. Higher education for sustainability: seeking affective learning autocomes. International Journal of Sustainability in Higher Education, vol 9 (1), 87-98, 2008.
31. Emmerling, R. J., Goleman, D. Emotional intelligence: Issues and common misunderstandings. *Issues and Recent Developments in Emotional Intelligence*,1(1), 2003.

32. Beard,C., Clegg,S. and Smith,K. (2007). Acknowledging the affective in higher education. British Educational Research Journal, 33, (2), 235-52, 2007.
33. Carlson, S. The sustainable university. The chronicle of Higher Education, October 20, 2006.
34. Adebayo, A. Student Engagement and Connection: Two Important Factors Often Not Emphasized In Teaching At The University Level. British Journal of Arts and Social Sciences. 4, (1), 2012.
35. Rowe, N., Wilkin, A., Wilson, R. Mapping of seminal reports on good teaching. (NFER Research Programme: Developing the Education Workforce). National Foundation for Educational Research, 2012.
36. Day, C., Sammons, P., Kington, A., Regan, E., Gunraj, J. and Towle, J. *Effective* Classroom Practice: a Mixed Method Study of Influences and Outcomes. Interim report submitted to the Economic and Social Research Council. Swindon: ESRC, 2008.
37. Dunne, M., Humphreys, S., Sebba, J., Dyson, A., Gallannaugh, F. and Muijs, D. Effective Teaching and Learning for Pupils in Low Attaining Groups (DCSF Research Report 011). London: DfES, 2007.
38. Siraj-Blatchford, I., Shepherd, D.L., Melhuish, E., Taggart, B., Sammons, P. and Sylva, K. Effective Primary Pedagogical Strategies in English and Mathematics in Key Stage 2: a Study of Year 5 Classroom Practice Drawn from the EPPSE 3-16 Longitudinal Study (DfE Research Report 129). London: DfE, 2011.
39. Garay,N., Cearreta,I.,López,J.M., Fajardo,I. (2006). Assistive technology and affective Mediation. Human Technology, vol 2(1) April,2006,pp 55-83.