2012 International Conference on Management and Education Innovation
IPEDR vol.37 (2012) © (2012) IACSIT Press, Singapore

Innovative Assessment Technologies: Comparing 'Face-to-Face' and Game-Based Development of Thinking Skills in Classroom Settings

Gyöngyvér Molnár [1] + and András Lőrincz [2]

[1] University of Szeged, [2] Eötvös Loránd University

Abstract. Technology offers new opportunities for education including educational games and complex methods for assessment. This paper aims at comparing the cognitive and affective results of a training program developed for first- and second-grade students in face-to-face and game-based environment and at investigating the opportunities of applying educational games and embedded innovative assessment technologies to improve students' reasoning skills. The training program consists of 120 educational games, developed directly for young learners. The experimental and control groups in the study consisted of 123 and 137 first and second graders, respectively. One third of the experimental group took part in a game-based training in computer-based environment (n=38), the others were involved in face-to-face training. Trainers in the face-to-face case asked relevant questions before and after the tests, whereas no dialogues were present in the game-based case. The training took six weeks to complete. The cognitive effectiveness of the training was measured with an inductive reasoning test, comprising 37 figural, non-verbal items (Cronbach α=.87). Besides the test-based data collection, facial expressions and head movement (captured by webcams) were monitored and the affective effects of the training were logged. The experimental group significantly outperformed the control group by more than one standard deviation on the post-test ($t=-10.65$, $p<.00$). The most frequent facial expressions were surprise (31%), happiness (24%), and anger (16%) during game-based training. Disgust (7%), fear (3%), and sadness (1%) were less frequent. The distributions of the facial expressions did not show significant relationships with the developmental level of reasoning skills. On the whole we found no significant differences between the developmental achievements of the face-to-face and game based groups after the training.

Keywords: Cognitive Development, Educational Games, Assessment, Emotion Monitoring

1. Introduction

Technology offers a great deal of new opportunities in education. It enhances the learning situation by allowing innovative task presentation, including multimedia, which increases motivation and the level of entertainment. Due to the technological development over the past decades a new training method has been gaining strength besides the traditional 'face-to-face' development programs: game-based development has become a viable option. Many studies argued that computer games should have a significant role in education (see Egenfeldt-Nielsen, 2007). Previous research showed that the development of abilities and reasoning skills via computer games is promising in many domains [Yelland, 2005]. However, there is a need for further research to support the application of educational theory to gaming since there are thoughtful warnings about the usefulness of technology in education [Toyama, 2011] and wise suggestions about the introduction of technology into schools [Cobo, 2011]. The design criteria for developmentally beneficial computer games must be based on widely-accepted education theories and practices [Peterson, Verenikina, & Herrington, 2008] and the effectiveness must be measured and compared to the traditional teaching methods.

Technology revolutionizes not only the learning situations but the embedded assessment and feedback mechanism as well [McGaw, Griffin, & Care, 2012]. Traditional methods provide the implementers with very few indicators, such as test scores, or subjective feedback regarding students' feelings during the

+ Corresponding author. Tel.: + 36 20 7756478; fax: +36 62 420034.
E-mail address: gymolnar@edpsy.u-szeged.hu.

training. Innovative assessment technologies provide innovative data entry and response capture, the teacher (and the computer) has access to rich data sets, which can lead to complex assessments methods by monitoring, logging, and analyzing emotional and cognitive issues in context [Csapó, Lőrincz, & Molnár, 2012].

In the present study, inductive reasoning strategies were trained explicitly face-to-face and in game-based environment using innovative assessment technologies for logging and analyzing metadata such as facial expressions based on the theory of basic emotions [Ekman, 1992]. The training is based on Klauer's theory of inductive reasoning and on his "Cognitive training for children" concept [Klauer & Phye, 2008].

Our main result is that there is no significant difference between training efficiency in face-to-face situation and in game-based environment for young pupils. This finding opens up a relatively inexpensive route for training inductive reasoning strategies for disadvantaged children at very early ages.

2. Objectives

Our motivation on developing high-quality low-cost training with remote assessment for many children gives rise to the following objectives: (1) compare the results and the effects of an inductive reasoning training program for first- and second-grade students in face-to-face and in game-based environments. (2) Investigate the opportunities for further studies that include both the cognitive and affective dimensions, especially with regards to the effectiveness of digital games as novel embedded innovative assessment technologies in order to improve students' reasoning skills.

In this paper we:

- explore the possibilities offered by technology-based assessment for a better understanding of students' inductive reasoning skills;
- compare the test-based results and the effects of an inductive reasoning training program for first and second graders in face-to-face and in game-based situations,
- present the results of the video-analyses on facial expressions that we detected (affective dimension);

3. Methods

3.1. Participants

The experimental group had 123 first- and second-grade students, whereas the control group consisted of 137 students having similar background variables. One third (38 students) of the experimental group took part in a game-based training in computer-based environment, while the remaining part participated in a face-to-face training.

3.2. Procedure

The training process was divided into 12 sessions, each lasting approximately 45 minutes. In the first and last sessions before and immediately after the training process test-based data collection took place to measure the cognitive dimension of the training-effect. The interval between the pre-test and the post-test was 8 weeks, the period in which the training was performed.

The program consists of 120 tasks (games) which can be solved through the application of appropriate inductive reasoning processes. In the face-to-face form of the training one quarter of the problems were manipulative requiring e.g. colorful building blocks, Dienes's logical set, matches, and so on [see Molnár, 2011]. The 120 computerized educational games were presented via touch screen monitors. Children used headsets and the training took place in a high-quality experimental classroom. Each game was started by a description of the task and had one or two additional questions that could be required by the children for additional help during the game.

In terms of the methods and work forms of the training, students were given the training individually. In the face-to-face environment there was a team of implementers, namely the class teachers, who introduced the activities to the children individually. In the computer-based training two facilitators were present in the classroom-like computer-lab during the whole training, where students were trained individually, however, at the same time in groups of 20.

The cognitive dimension of the training-effect was measured by traditional test-based data collection in both situations. The test comprised 37 figural, non-verbal items (Cronbach $\alpha = .87$). Test construction and a precise alignment between the framework and the test ensured validity of the test [see Molnár, 2011].

Besides the test-based data collection, innovative assessment technologies were used in the game-based environment with the intent of monitoring the affective processes, e.g. logging and analyzing metadata, such as head movement and facial expressions. Webcam-based technology and off-line facial expression evaluation was utilized [Jeni et al. 2012]. However, the technology offers real time evaluation and real time feedback [Saragih et al., 2011]. The method is sketched in Figure 1a-b. First, for each landmark, the response map – a probability distribution – is computed (Figure 1a). Then the results, i.e., the individual probability distributions are combined to form the best probability estimation subject to the constraints of the form of the face. Further improvements are achieved step-by-step by sliding towards the minimum along the negative gradient of the objective function of the so called constrained local model (CLM) [Cristinacce & Cootes, 2008] that we applied. The result of this successive approximation gives rise to the estimation of the landmark positions. We have estimated the six basic emotions, namely happiness, sadness, disgust, fear, anger, and surprise (Figure 1c). Our multi-class classifier system was trained on the CK+ [Lucey, 2008] and on the BU 4DFE [Yin, 2008] databases. Recognition reached about 87% (for more details, see [Jeni, 2012]).

(a) (b) (c)

Fig. 1.: (a): Response map for the right corner of the right eye. (b): estimated positions of the relevant landmarks on the face for emotion estimation, (c): Left hand side: original image; right hand side: results of the evaluation. Crosses: landmark points of the automated annotation. Roll, yaw, pitch: angle of head pose. Anger, etc: emotions. Rectangles: individual estimations of the different emotion sensors that may be misguided by similarities of the different emotions. Larger open rectangle: classification of the emotion.

The affective dimension of the training cannot be measured in the face-to-face environment presently. Lacking the objective measurement tools, only the subjective opinion of the implementers gave us some information regarding the affective processes. However, future developments will provide such objective tools since our methods are pose independent [Jeni, 2012] and may reach the sufficient level of robustness in the near future.

4. Results

4.1. Cognitive dimension of the training effect

No significant differences were found between the performance of the experimental and the control groups in face-to-face ($t=1.06$, $p>.05$) and in game-based ($t=1.44$, $p>.05$) environment prior to the experiment, while the experimental groups managed to achieve significant development in the experimental period in both cases ($t=15.42$, $p<.001$; $t=-10.65$, $p<.001$), respectively. The experimental groups significantly outperformed the control group by more than one standard deviation on the post-test (Figure 3) and nearly each student attained significant improvement in their performance as a result of the training independent from the used media.

There were no significant differences between the post-test mean scores of students getting the training in face-to-face or game-based environment ($t=1.70$ $p>.05$). Nevertheless, a different tendency can be observed regarding the performance of the control group. Some students underperformed their pre-test performance by almost 30% in the post-test. The above results are supported by the student level analyses. There were no students in the experimental group whose performance dropped significantly from pre-test to post-test; moreover, several students improved in both modalities by more than one standard deviation. The effect size of the face-to-face training program was $d=1.05$ ($p<.01$) and it was $d=.87$ ($p<.01$) regarding the game-based environment. Using Cohen's convention for describing the magnitude effect size, these are clearly large effects.

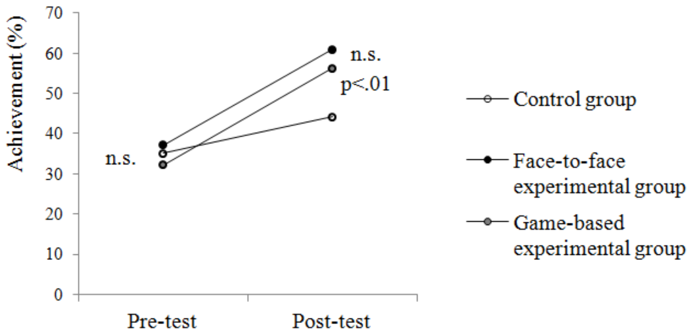

Fig. 3.: Changes of the mean achievement of the control and the face-to-face and game-based experimental group from pre-test to post-test

4.2. Affective dimension of the training effect

The precision of our emotion estimation system was tested on the FaceGen [Jeni, 2011] and on the BU 4DFE [Jeni, 2012] databases. State-of-the-art emotion recognition was achieved and the system showed robustness against pose variations that – together – enable emotion estimation in the classroom.

In face-to-face environment it was not possible to do any objective, comparable analyses because of the subjectivity of the implementers introduced the activities to the children. In game-based environment the most frequent facial expressions were those of surprise (31%), happiness (24%) and anger (16%) during the game-based training. Disgust (7%), fear (3%) and sadness (1%) were less frequent. Children were not always facing the screens, therefore, there was no detectable face on 19% of the frames. No significant differences were found between genders regarding the distribution of the different expressed emotions. There were no significant correlations between the level of the cognitive development and the distributions of the facial expressions.

5. Discussion

This paper addresses the comparison of the cognitive and affective effectiveness of a training program for Grade 1 and 2 students in face-to-face and game-based environment, presents the underlying possibilities for a computer based intervention and presents the direct result of the evaluation study. According to the analyses the developmental level of students in the experimental and control group did not differ prior to the experiment and similarly mean test scores of students getting the training face-to-face or game-based did not differ either prior to the experiment. It means that the two subsample of the experimental group meets the requirement for doing the comparisons within the two subsample of the experimental group in the study as well.

As a cognitive result of the training, the reasoning skills of the experimental group students showed greater improvement than that of the control group students in both environments. Thus, our findings provided empirical evidence that computer games even without intelligent tutor are also efficient tools in the development of reasoning skills at this early age.

Regarding the cognitive effectiveness of the training the comparison of the results on student level made it possible to consider the changes in face-to-face and game-based experimental environment and to compare the changes in control group performance based on students' original level of inductive reasoning skills as well. The performance pattern of the students in the experimental and control group differ significantly, while similar tendency is displayed in connection with the fact-to-face and game-based experimental group students' reasoning skills. There were no students in the experimental group independent from the used media whose performance dropped significantly from pre-test to post-test. Most of the students' performance in the face-to-face and game-based environment attained improvement by more than one standard deviation as a result of the training. The training regardless of the used media enhanced students' reasoning with highly diverse levels of skills. The effect sizes achieved in face-to-face and game based environment are outstanding internationally as well.

We have outlined how contextual information can be gathered while students play the game. According to the facial expression-analyses the two emotions which we can consider positive (happiness, surprise) have been experienced more frequently during the training than anger, disgust, fear or sadness. This suggests that using our game-based program provides children with enjoyable time. Having the emotion of surprise in this large quantity imply that our program does not seem to be unexciting or uninteresting for the users.

The results suggest that game-based environment do not decrease the effectiveness of an educational training program, it can be used in a computer-based environment with similar level of effectiveness as one-to-one development in face-to-face environment. With sufficient educational game design the advantages of the game-based environment are clear: the presence of the class school teacher is not crucial during the development process and students can be trained in groups. These characteristics of the game-based media indicate great potential for educational use.

6. Acknowledgements

The present study was conducted with the financial support of the K75274 OTKA research program, the Szeged Center for Research on Learning and Instruction and the SZTE MTA Research Group on the Development of Competencies. Gyongyver Molnar was on a Bolyai János Research Fellowship at the time the present study was written. The research was also funded by the European Union, co-funded by the European Social Fund and the European Regional Development Fund (Project ID numbers: TÁMOP 3.1.9-08/1-2009-0001, TÁMOP 4.2.1/B-09/KMR-2010-003, KMOP-1.1.2-08/1-2008-0002).

7. REFERENCES

[1] C. Cobo. How we use technology in education is more important than which technology we use. EduTechDebate 2011, https://edutechdebate.org/ict-in-schools/there-are-no-technology-shortcuts-to-good-education/

[2] D. Cristinacce, and T. Cootes. Automatic feature localization with CLMs, *Pattern Rec.* 2008, **41**: 3054–3067.

[3] B. Csapó, A. Lőrincz and G. Molnár. Innovative Assessment Technologies in Educational Games Designed for Young Students. In: D. Ifenthaler, D. Eseryel, and X. Ge (eds.). *Assessment in game-based learning: foundations, innovations, and perspectives*. New York: Springer. in press. 2012.

[4] S. Egenfeldt-Nielsen. Making sweet music: The Educational Use of Computer Games. In: J. Linderoth (ed.) *Datorspelandets Dynamik. Studentlitteratur*. 2007. http://egenfeldt.eu/papers/sweet_music.pdf

[5] P. Ekman. An argument for basic emotions. *Cognition & Emotion*, 1992, **6**: 169-75, 192-200.

[6] L. A. Jeni, H. Hashimoto, and A. Lőrincz. Pose Invariant Facial Expression Classification based on HighQuality Shape Information. 2011. IEEE CVPR Workshop on Gesture Recognition, Colorado Springs, CO, USA June 20.

[7] L. A. Jeni, A. Lőrincz, T. Nagy, Z. Palotai, J. Sebők, Z. Szabó, and D. Takács. 3D shape estimation in video sequences provides high precision evaluation of facial expressions. Image and Vision Computing (in press) 2012.

[8] K. J. Klauer, and G. D. Phye. Inductive reasoning. *A training approach. Rev. Edu. Res.*, 2008, **78**: 85-123.

[9] P. Lucey, J. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews. The extended Cohn-Kanade dataset (CK+). In: *3rd IEEE Workshop on Human Communicative Behavior Analysis (CVPR4HB10)*. 2010, pp. 94–101.

[10] B. McGaw, P. Griffin, and E. Care (eds.). *Assessment and teaching of 21st century skills*. NY Springer. 2012.

[11] G. Molnár. Playful fostering of 6- to 8-year-old students' inductive reasoning. *Thinking skills and Creativity*, 2011, **6** (2): 91-99.

[12] R. Peterson, I. Verenikina, and J. Herrington. Standards for educational, edutainment, and developmentally beneficial computer games. In: J. Luca and E. Weippl (Eds.). *Proc. of EdMedia*, AACE, Chesapeake, VA, 2008, pp. 4143-4151.

[13] J. M. Saragih, S. Lucey, and J. F. Cohn. Deformable model fitting by regularized landmark mean-shift. *Int. J. Comp. Vis.*, 2011, **91**: 200–215.

[14] K. Toyama. There are no technology shortcuts to good education. ICT in Schools, EduTechDebate. 2011. https://edutechdebate.org/ict-in-schools/there-are-no-technology-shortcuts-to-good-education/

[15] N. Yelland. The future is now: A review of the literature on the use of computers in early childhood education (1994–2004). *AACE Journal*, 2005, **13** (3): 201-232.

[16] L. Yin, X. Chen, Y. Sun, T. Worm, and M. Reale. A high-resolution 3D dynamic facial expression database. In: *8th IEEE International Conference on Automatic Face Gesture Recognition*, IEEE, 2008, pp. 1–6, Number: 66.