João Dias
Pedro A. Santos
Remco C. Veltkamp (Eds.)

LNCS 10653

Games and Learning Alliance

6th International Conference, GALA 2017
Lisbon, Portugal, December 5–7, 2017
Proceedings

Springer

Lecture Notes in Computer Science 10653

More information about this series at http://www.springer.com/series/7409

João Dias · Pedro A. Santos
Remco C. Veltkamp (Eds.)

Games and Learning Alliance

6th International Conference, GALA 2017
Lisbon, Portugal, December 5–7, 2017
Proceedings

Springer

Editors
João Dias
University of Lisbon
Lisbon
Portugal

Pedro A. Santos
University of Lisbon
Lisbon
Portugal

Remco C. Veltkamp
Utrecht University
Utrecht
The Netherlands

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-319-71939-9 ISBN 978-3-319-71940-5 (eBook)
https://doi.org/10.1007/978-3-319-71940-5

Library of Congress Control Number: 2017959629

LNCS Sublibrary: SL3 – Information Systems and Applications, incl. Internet/Web, and HCI

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The 6th Games and Learning Alliance (GALA) Conference, organized by the Serious Games Society (SGS) and the Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa (INESC-ID), was held in Lisbon, Portugal, December 5–7, 2017.

The GALA series of conferences provides an excellent opportunity to foster the discussion of important themes and topics in the growing field of serious games. The conference is a venue for academic researchers, industrial developers, teachers, and corporate decision makers to meet and exchange experiences and knowledge in this multidisciplinary and challenging area.

GALA 2017 received 45 submissions. Each paper was reviewed by at least three Program Committee members. The Program Committee selected 24 of these papers for presentation at the conference, and 9 for presentation at a poster session of the conference. A total of 17 countries were represented at the conference. While the majority of authors are based in Europe, North and South America and Asia were also represented at the conference.

The conference started with a day of tutorials, focusing on two main areas and their application in the context of serious games: the development of socio-affective characters for interacting with users, and data analytics and machine learning techniques. The European project RAGE contributed significantly to these tutorials.

It was an honor to have Ana Paiva, from the Instituto Superior Técnico, Universidade de Lisboa, Portugal, as keynote speaker at GALA 2017. Ana Paiva discussed the use and role of social robots in the creation of novel and interesting serious games.

The conference featured six paper presentation sessions. A number of paper presentation sessions discussed the relation between serious games and particular domains, namely, children with disabilities, management, persuasive games, augmented reality, game analytics, game design, health, maths, MOOCs, gamification, and virtual reality. Other sessions discussed game development and the assessment of games, and the relation between games and learning.

As in previous years, selected best papers of the GALA conference will be published in a dedicated special issue of the International Journal of Serious Games, the scientific journal managed by the Serious Games Society, which is a great reference point for academicians and practitioners to publish original research work on serious games and be informed about the latest developments in the field.

We thank the authors for submitting many interesting papers, the Program Committee for reviewing these papers, and the SGS and INESC-ID for organizing the conference.

October 2017

João Dias
Pedro A. Santos
Remco C. Veltkamp

Organization

GALA 2017 was organized by the Serious Games Society and the Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa (INESC-ID).

General Chair

João Dias	INESC-ID and Instituto Superior Técnico, Universidade de Lisboa, Portugal

Program Chairs

Pedro A. Santos	INESC-ID and Instituto Superior Técnico, Universidade de Lisboa, Portugal
Remco C. Veltkamp	Utrecht University, The Netherlands

Program Committee

Anissa All	University of Gent, Belgium
Alessandra Antonaci	ITD-CNR, Italy
Sylvester Arnab	Coventry University, UK
Jannicke Baalsrud Hauge	BIBA, Germany/KTH, Sweden
Per Backlund	Högskolan i Skövde, Sweden
Norman Badler	University of Pennsilvania, USA
Sylvie Barma	Université Laval, Canada
Francesco Bellotti	University of Genoa, Italy
Riccardo Berta	University of Genoa, Italy
Rafael Bidarra	TU Delft, The Netherlands
Staffan Björk	Chalmers, Sweden
Rosa Bottino	ITD-CNR, Italy
Clint Bowers	University of Central Florida, USA
Maira Brandao Carvalho	Tilburg University, The Netherlands
Sylvie Daniel	University Laval, Canada
Alessandro De Gloria	University of Genoa, Italy
Kurt Debattista	University of Warwick, UK
Shujie Deng	University of Bournemouth, UK
Michael Derntl	RWTH Aachen, Germany
Frank Dignum	University of Utrecht, The Netherlands
Ioannis Doumanis	University of Middlesex, UK
Miguel Encarnação	University of Louisville, USA
Manuel Gentile	ITD-CNR, Italy

Junghyun Han	Korea University, South Korea
Valerie Hill	Texas Woman's University, USA
Carolina Islas Sedano	University of Eastern Finland, Finland
Michael Kickmeier-Rust	University of Graz, Austria
Sotiris Kriginas	University of Athens, Greece
Ralph Klamma	RWTH Aachen, Germany
Silvia Kober	University of Graz, Austria
Niki Lambropoulos	Global Operations Division, Greece
George Lepouras	University of Peloponnese, Greece
Theo Lim	Heriot-Watt University, UK
Sandy Louchart	Glasgow School of Art Digital Design Studio, UK
Katerina Mania	Technical University of Crete, Greece
Fabrizia Mantovani	Università di Milano Bicocca, Italy
Michela Mortara	IMATI-CNR, Italy
Rob Nadolski	Open University, The Netherlands
Manuel Ninaus	Leibniz-Institut für Wissensmedien, Germany
Jussi Okkonen	University of Tampere, Finland
Magda Popescu	Carol I Nat. Defense University, Romania
Samuel Mascarenhas	INESC-ID, Lisbon, Portugal
Marius Preda	Institut National des Télécommunications, France
Matthias Rauterberg	TU Eindhoven, The Netherlands
Ion Roceanu	Carol I Nat. Defence University, Romania
Margarida Romero	Université Nice Sophia Antipolis, France
Esther Judith Schek	University of Milan-Bicocca, Italy
Avo Schönbohm	Berlin School of Economics and Law, Germany
Ioana Stanescu	Carol I Nat. Defence University, Romania
Kam Star	Playgen, UK
Erik van der Spek	TU Eindhoven, The Netherlands
Herre van Oostendorp	Utrecht University, The Netherlands
Remco Veltkamp	Utrecht University, The Netherlands
Wim Westera	Open University, The Netherlands
Antonie Wiedemann	University of Genoa, Italy
Peter Wolf	ETH Zuerich, Switzerland
Josef Wolfartsberger	University of Applied Sciences Upper Austria, Austria
Su Ting Yong	University of Nottingham Malaysia Campus, Malaysia
Zerrin Yumak	Utrecht University, The Netherlands

Awards Chair

Jannicke Baalsrud Hauge	BIBA, Germany/KTH, Sweden

Awards Jury

Sylvester Arnab	Coventry University, UK
Massimiliano Cazzaniga	i-maginary, Italy
George Lepouras	University of Peloponnese, Greece

Margarida Romero	Université Nice Sophia Antipolis, France
Harald Warmelink	Breda University of Applied Sciences, The Netherlands
Pau Yanez	Geomotiongames, Spain

Demo Chair

Giorgio Da Bormida	ELGI Consulting, Inspire, Clear Communication Associates Ltd.

Tutorial Chair

Samuel Mascarenhas	INESC-ID, Lisbon, Portugal

Publications Chair

Riccardo Berta	University of Genoa, Italy

Communication and Promotion Chair

Francesco Bellotti	University of Genoa, Italy

Administrative and Financial Chair

Antonie Wiedemann	University of Genoa, Italy

Contents

Gamification

Augmented and Virtual Reality

Methods and Tools (for Design and Development)

Poster Abstracts

Games in Education and Training

Data-Driven Design Decisions to Improve Game-Based Learning of Fractions

Manuel Ninaus[1,2](✉), Kristian Kiili[3], Robert S. Siegler[4,5], and Korbinian Moeller[1,2,6]

[1] Leibniz-Institut für Wissensmedien, Tuebingen, Germany
{m.ninaus,k.moeller}@iwm-tuebingen.de
[2] LEAD Graduate School, Eberhard-Karls University, Tuebingen, Germany
[3] TUT Game Lab, Tampere University of Technology, Pori, Finland
kristian.kiili@tut.fi
[4] Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA
rs7k@andrew.cmu.edu
[5] Siegler Center for Innovative Learning, Beijing Normal University, Beijing, China
[6] Department of Psychology, Eberhard Karls University, Tuebingen, Germany

Abstract. The educational value of games for learning is of prime importance. To design tasks that promote learning, findings from basic research and user data from earlier versions of the game can be very helpful. In the current study, we reanalysed data from two studies using a game-based assessment and a training tool for improving fraction magnitude knowledge. Similar to results from basic research, data of both studies indicated that performance in game-based versions of number line estimation and magnitude comparison were highly correlated (even after controlling for students' overall math achievement). This suggests that both tasks share an underlying (spatial) representation of number magnitude. Based on this we designed a new task to improve fraction learning. In particular, we merged mechanics of number line estimation and magnitude comparison to consider shared processing of number magnitude in both tasks in our game-based learning of fractions.

1 Introduction

The use of data from playtesting earlier versions of a game is a common and valuable approach to improve user experience of entertainment games (for a basic introduction see [1]). These so-called game analytics aim at discovering and communicating patterns in data to inform, for instance, the ongoing design process of a game to optimize user experience. Such game analytics might be even more important in developing successful educational games by not only optimizing user experience but learning outcomes in particular (e.g., [2, 3]). In the current study, we discuss the background of developing a new task and game mechanic based on previous research results and reanalyses of user data from an earlier version of an educational game designed for learning rational numbers.

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 3–13, 2017.
https://doi.org/10.1007/978-3-319-71940-5_1

Educational games are primarily defined by their educational value. While general models such as the Learning Mechanics-Game Mechanics (LM-GM) model [4] can support the development process, a more in-depth and domain specific analysis is necessary to optimize learning outcomes. As such, findings from basic research can inform design decisions of an educational game early on (e.g., [5–7]). During the development process of the game employed in the current study, we considered recent findings from numerical cognition research (e.g., [8] for a review on rational number knowledge) to adapt game mechanics for increasing learning outcomes.

The game of interest, *Semideus*, primarily focuses on assessing and improving rational number (i.e., fractions and decimals) knowledge in young students. Fractions are one of the most challenging topics in mathematics education [9]. They are also crucial for learning of more advanced mathematics, as indicated by it being a very good predictor of future algebra performance and overall mathematics achievement [10, 11]. Unfortunately, many students fail to master fractions [8, 9]. One major difficulty for students is understanding fraction magnitude (for a review, see [8]). Importantly, two types of tasks are used primarily for assessing and improving fraction magnitude knowledge: *number line estimation* and *magnitude comparison* tasks (e.g., [12–15] for a recent review see [16]). Consequently, these two tasks were considered crucial for a game, which focuses on improving rational number knowledge.

The *number line estimation task* is based on the metaphor of a mental number line, according to which small/large magnitudes are associated with the left/right side of space in western cultures (for an overview see [17]). Accordingly, in number line estimation (e.g., [18]), participants have to indicate the spatial position of a target number on a horizontal line with only its start and endpoint specified [e.g., where does $4/5\,(67)$ go on a number line ranging from 0–1 (0–100)]. Basic research suggests that training of such spatial representations of number magnitude corroborates more than just the accurate mapping of numbers on the mental number line but generalizes to other numerical competencies. As a consequence, training involving the concept of a mental number line seems to be particularly effective when promoting young students' numerical competencies (e.g., [19, 20]). Hence, *Semideus* uses number line estimation as its most basic task.

In the typical *number magnitude comparison task*, participants have to decide which of two numbers is larger [e.g., $4/5$ (80) is larger than $2/3$ (66)]. Interestingly, participants' responses get longer and are more error prone the closer in magnitude the to-be-compared numbers get (e.g., [21, 22]). This so-called numerical distance effect indicates a successful spatial representation of number magnitude and supports the idea of a mental number line even though, actual comparison of spatial magnitudes is not explicitly required in this task. Interestingly, performance on both *number line estimation* and *number magnitude comparison task* using whole numbers (e.g., [23–25]) is highly correlated for both whole numbers and fractions, though so far only two studies have examined this relation with fractions [12, 26]. Importantly, studies involving training of number line estimation often use magnitude comparison as an evaluation task to assess training effects (e.g., [27]). It is assumed that it should be easier to differentiate numbers by their magnitude after participants have been trained to estimate the location of a magnitude on a number line. However, at least for fraction learning, it is unclear whether this relation is preserved when controlling for individuals' overall math achievement or math grades, respectively, which hasn't been done in previous studies. In case this relationship disappears, one might argue that the association between number line estimation and magnitude comparison is due to better overall math performance rather than indicating a shared underlying representation of number magnitude.

When implementing magnitude comparison in our game, we changed the conventional number magnitude comparison task a little bit to enhance students' spatial representation of number magnitude. In particular, we designed the comparison to take place on a number line with its endpoints defined as "small" on the left and "large" on the right (see Game Description and Fig. 1 right chart). Participants had to place numbers on the number line in correct ascending order. On this task, absolute correct positioning of to-be-compared numbers is irrelevant, but the basic characteristics of a comparison task are maintained, and spatial aspects of number magnitude might be enhanced. However, in a recent training study [15], we only identified significant improvements on estimation tasks when employing rational number estimation and our adapted comparison task. We argued that the magnitude comparison task in its current form might not accurately assess the spatial representation of number magnitude. Improving this magnitude comparison task might be necessary to accurately assess rational number knowledge.

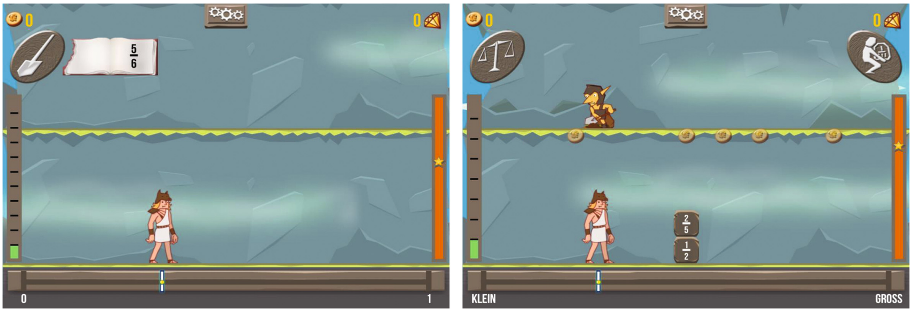


Fig. 1. Left chart: Example of an estimation task; Right chart: Example of a comparison task; klein = small, gross = large.

To increase the educational value of our game, the current study examined and reanalysed data from two previous studies [14, 15] employing our game-based estimation and adapted magnitude comparison task in two contexts (i.e., assessment and training) and two age groups (4th and 5th graders). We evaluated the relation between performance in number line estimation and magnitude comparison. The goal was to examine whether these two tasks share the same underlying representation. This would indicate that a combination of number line estimation and magnitude comparison task mechanics might improve future assessment and training of rational number knowledge.

2 Method

In Study 1 ([14], Assessment-study), our game-based learning environment was used to assess conceptual knowledge of fractions in Finnish fifth graders. We were able to replicate hallmark effects of fraction magnitude processing typically observed in basic research, such as the numerical distance effect. This suggested that game-based learning environments for fraction education may also allow for a valid assessment of students' fraction knowledge. In Study 2 ([15], Training-study), we employed the same game-based environment as a training tool to improve fourth graders conceptual knowledge of fractions. Results indicated that the game-based training group improved their conceptual knowledge of fractions more strongly than a control group. Reanalysing data from both studies allowed us to examine and compare data from different contexts and age groups. This provides us with the promising opportunity to revisit earlier iterations from a game-based learning environment to influence the ongoing design process. Importantly, analysing and comparing data from a training and assessment study allows us to transfer our results from one context to the other. In this section, we first describe the game based environment and then other relevant aspects of both studies. For a more comprehensive description of the two studies see [14, 15].

2.1 Game Description

In both studies, similar versions of the game-based learning application *Semideus* were used. *Semideus* is a game engine that allows for creating game-based tasks for improving and assessing rational number knowledge (e.g., [14, 28]). The game is set in ancient Greek times. Users control the avatar "Semideus" (by tilting the tablet), who tries to find and retrieve gold coins that the goblin Kobalos has stolen from Zeus. Kobalos has hidden the coins along the trails of Mount Olympos. Semideus has to discover the locations of the hidden coins, encrypted in mathematical symbols (e.g., fractions). The core gameplay requires working with number lines (e.g., [26]) to perform number line estimations and magnitude comparisons. Number lines were implemented as walkable platforms.

In the estimation tasks, users had to move the avatar to the correct position on the number line. Figure 1 (left chart) shows an example of an estimation task in which players had to locate the fraction 5/6 on a number line. Estimates more than 8% off from the correct position were categorized as inaccurate. For correct estimates, players received 100–500 coins depending on the degree of accuracy.

In our adapted magnitude comparison task, users had to arrange two stones with an engraved number on them in ascending order on a number line ranging from "small" to "large". Hence, the absolute position of to-be-compared numbers on the number line was irrelevant as long as the relative position of the stones to each other was correct. For instance, players had to place the stone "2/5" left from stone "1/2" for the placement to be correct. Correct comparisons were rewarded with 100-500 coins depending on the response time of the users, with higher rewards for faster responses.

2.2 Study 1 (Assessment Study)

Participants: Fifty-four fifth-graders (25 male; mean age = 11.26 years, SD = 0.48 years) participated in the study. They were equipped with iPads and had 30 min to complete the game. Seven students failed to provide their math grades. Thus, only 47 students were included in the final analyses.

Procedure: All students were examined during regular school hours. First, experimenters introduced the game and explained game mechanics to students. Then, students received their user account, which was used to record individual game behaviour. Each student received an iPad and played the game individually. They were not allowed to discuss the game with other students during the play session. Math achievement was measured by participants' previous math grade (Finnish classification scheme: 10 reflects the highest and 4 the lowest grade).

2.3 Study 2 (Training Study)

Participants: In the training group, 68 fourth-graders played the game, of who 54 (mean age = 10.24 years; SD = .43; 25 males) followed the requested protocol. That is, they participated in both the pre- and posttest and played the game in-between. Another 45 students were recruited for the control group, of who 41 (mean age = 10.02 years; SD = .27; 25 males) participated in both the pre- and posttest. These children did not play the game and therefore are not considered here.

Procedure: Experimenters explained the game to students before they started to play it. Students had to play the game in 5 sessions of about 30 min each during a four-week period. There was no additional teaching of rational numbers in school during the study period. Math achievement was again measured by participants' previous math grade in the Finnish classification system. Log files of students' estimation and comparison performance during the games were analysed.

2.4 Analysis

Correlations were computed between fraction magnitude comparison performance and fraction estimation accuracy for fractions between 0 and 1 for both the Assessment (Study 1) and the Training study (Study 2). We also conducted partial correlation analyses between fraction magnitude comparison performance and fraction estimation accuracy, controlling for the effects of students' previous math grade (overall math performance). The control variable (math grade) is the variable which extracts the variance which is obtained from the initial correlation between fraction magnitude comparison performance and fraction estimation accuracy. This allowed us to investigate whether the correlation between performance in these two tasks is still present when controlling for overall math performance. The correlations and partial correlations were conducted using R [29] and the R package corrplot [30]. Data visualization was realized with the package ggplot2 [31].

3 Results

3.1 Study 1 (Assessment Study)

Students' performance in fraction comparison and fraction estimation was positively correlated [r(45) = 0.64, p < .001, see also Fig. 2 Panel A]. Moreover, after controlling for individual's overall math achievement the correlation remained positive and significant [r(44) = . 41, p < .005, see also Fig. 2 Panel B].

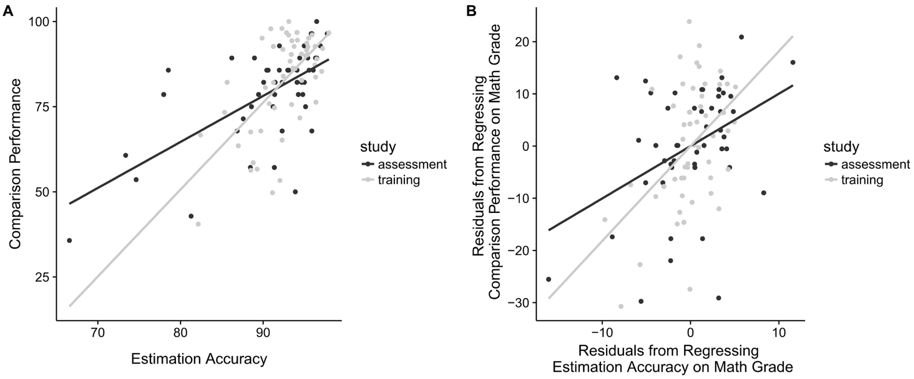


Fig. 2. Panel A: Scatterplot for comparison performance and estimation accuracy; Panel B: Scatterplot for residuals of comparison performance and estimation accuracy controlling for math grade.

3.2 Study 2 (Training Study)

As in Study 1, students' performance in fraction comparison and fraction estimation was significantly correlated [r(52) = .62, p < .001, see also Fig. 2 Panel A]. Again, the correlation remained significant even after controlling for individual's overall math achievement [r(51) = .46, p < .001, see also Fig. 2 Panel B].

4 Discussion

In the present study, we reanalysed data from two previous studies [14, 15] to better understand the relation between number line estimation and magnitude comparison. Medium to large correlations remained between estimation accuracy and comparison performance of fractions even when controlling for overall math achievement, suggesting that both tasks draw on the same underlying representation of number magnitude. In the following, we consider the results of the current study and recent literature to design and propose a new way of assessing and training conceptual knowledge of (rational) number magnitude. This data-driven approach may result in new tasks/mechanics, which might be more beneficial not only for learners but also for educators and researchers in the domain of numerical cognition, as they might provide more detailed information about users' strategies and competencies.

Current and previous studies indicate that number line estimation and magnitude comparison share the same underlying representation of number magnitude [20, 24, 25]. Our data show, that this cannot be explained by individuals' overall math performance, as we controlled for this. However, this also suggests that previous implementations of the comparison task within *Semideus* may not have been optimal as we did not observe significant improvements in comparison performance following the training procedure [15]. This was surprising, because correlations between number line estimation and magnitude comparison performance in fractions were already described elsewhere [12, 26]. To overcome this shortcoming of our current implementation we redesigned the comparison task mechanics. In order to draw users' attention more explicitly to the use of spatial locations on the number line we combined mechanics of both number line estimation and magnitude comparison. By merging these two tasks, we aim to further foster fraction magnitude understanding in future training studies. In the following, we elucidate this data-driven design decision by detailing the new mechanic:

In its current version, *Semideus*, as a game-based approach, allows us to seamlessly integrate magnitude comparisons and number line estimates into the gameplay by narrative elements of the game. In particular, the main task of users is to dig up a gold coin at a location on the number line specified at the top left corner besides the shovel button (target number; see Fig. 1 left chart). This reflects the basic mechanic of a number line estimation task (e.g., [18]). By integrating "traps" (i.e., positions to avoid) on the number line, it is possible to integrate number magnitude comparisons into the gameplay. The location of the traps is defined in the top right corner (a lightning symbol refers to traps; see Fig. 3 A–D). Accordingly, users need to decide whether a trap is on their way to the position they need to walk to in the primary number line estimation task. This decision reflects number magnitude comparison (see Fig. 3 A: 2/3 is larger than 4/9; Fig. 3 B: 3/7 is smaller than 2/3). If a user walks through a trap, he/she loses virtual energy (see Fig. 3 C).

To integrate magnitude comparison into the narrative of the gameplay, users are asked to disarm the trap when it is on his/her way to the position of the gold coin. Players can disarm the trap by pressing a button located on top of the screen (symbol of a mole; see Fig. 3 A–D). When the user has disarmed the trap with the mole, he/she can walk over it and dig up the coin at the estimated position. To decide whether disarming the trap is necessary, users have to compare the location of the coin (target number) with the location of the trap and based on this decision disarm or not disarm the trap. This requires an explicit comparison of the magnitudes of the target number and the one specifying the location of the trap. Accordingly, this should increase the association of fraction magnitudes with spatial locations on the number line (see a demo of this idea on youtube: https://youtu.be/cFS7USJJ3pI).

Moreover, this suggested comparison mechanic might make comparison tasks easier to solve. The previous comparison mechanic turned out to be challenging for some students. More specifically, some users had difficulty adopting game controls needed to perform the comparison [32].

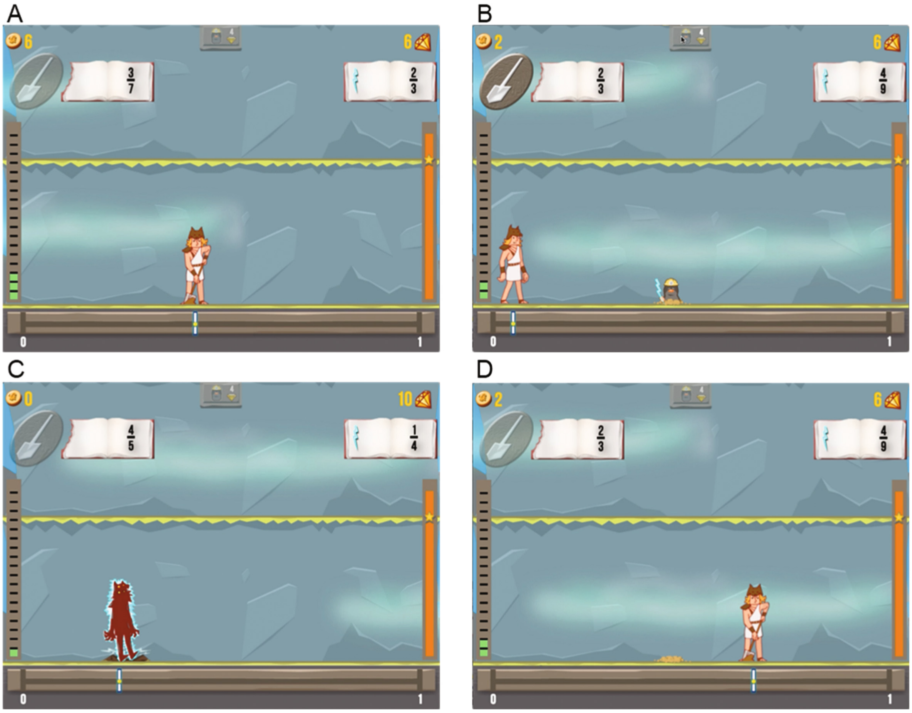


Fig. 3. Combination of number estimation and magnitude comparison task mechanics; target location/number defined in the top left corner; traps/locations to avoid defined in the top right corner with a lightning symbol; A: Target number is smaller than trap location; user can safely dig out coins without disarming the trap. B: Target number is larger than trap location; user needs to use the mole to disarm the trap; C: User receives negative feedback when walking over a trap and looses virtual energy (orange bar on the right of the screen; D: User has used the mole to disarm the bomb (mound of earth) to safely walk to the target location.

A further iteration of the current task mechanic, which is being developed, addresses adaptive gameplay, this means user-specific adjustments of task difficulty or game progress. Specifically, instead of a mole disarming the trap at a given position on the number line, users may use a piece of wood as a bridge to cross the trap without harm. This alternative mechanic will have pieces of wood of different lengths. In the beginning, large pieces of wood would cover a wider range of the number line. Thus, knowing approximately where the trap is would be sufficient to avoid it. The available, pieces of wood will become smaller as the person becomes more accurate and advances in the game, which requires more specific localisation of traps. This addition to the task mechanic provides a natural way of changing task difficulty and might also inform researchers and educators about students' strategies when solving such tasks.

Future work: Future studies need to evaluate whether this new task/mechanics might (i) be superior to conventional training and assessment methodologies in producing and measuring learning, (ii) provide more information about students' strategies in solving such fraction magnitude tasks, (iii) extend task diversity, (iv) improve user experience, and (iv) add an additional layer of difficulty for already well performing students. Moreover, in the future we aim at employing similar data and theory-driven approaches to evaluate new design iterations of the game. This will also include the use of game analytics tools to investigate not only performance but also player experience variables, which might also allow us to identify more complex patterns in the data.

Conclusion: By reanalysing data from an educational game for learning fractions, we identified an association between number line estimation and number magnitude comparison, even after grades in math courses were statistically controlled. This suggests that both tasks share the same underlying representation of number magnitude. Based on these results and those in previous studies in the domain of numerical cognition, we designed a new task combining task mechanics of number line estimation and number magnitude comparison by including the need for explicit magnitude comparison in our number line based game. This new task might improve future training of rational number knowledge. Most important, although the current study described task mechanics for an educational game, its use is not limited to gaming contexts. Instead, the technology might as well be utilized in a wide range of areas of numerical cognition.

Acknowledgments. The current research was supported by the Leibniz-Competition Fund (SAW) supporting Manuel Ninaus (SAW-2016-IWM-3) and the Leibniz-WissenschaftsCampus "Cognitive Interfaces" and Academy of Finland (grant number 289140) supporting Kristian Kiili.

References

1. Drachen, A., Seif El-Nasr, M., Canossa, A.: Game Analytics – The Basics. In: El-nasr, M.S., Drachen, A., Canossa, A. (eds.) Game Analytics, pp. 13–40. Springer, London, London (2013). https://doi.org/10.1007/978-1-4471-4769-5_2
2. Freire, M., Serrano-laguna, Á., Iglesias, B.M., Martínez-Ortiz, I., Moreno-Ger, P., Fernandez-Manjon, B.: Game learning analytics: learning analytics for serious games. In: Spector, M.J., Lockee, B.B., Childress, M.D. (eds.) Learning, Design, and Technology, pp. 1–29. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-17727-4_21-1
3. Hauge, J.B., Berta, R., Fiucci, G., Manjon, B.F., Padron-Napoles, C., Westra, W., Nadolski, R.: Implications of Learning Analytics for Serious Game Design. In: 2014 IEEE 14th International Conference on Advanced Learning Technologies, pp. 230–232. IEEE (2014)
4. Arnab, S., Lim, T., Carvalho, M.B., Bellotti, F., de Freitas, S., Louchart, S., Suttie, N., Berta, R., De Gloria, A.: Mapping learning and game mechanics for serious games analysis. Br. J. Educ. Technol. **46**, 391–411 (2015)
5. Räsänen, P., Salminen, J., Wilson, A.J., Aunio, P., Dehaene, S.: Computer-assisted intervention for children with low numeracy skills. Cogn. Dev. **24**, 450–472 (2009)
6. Käser, T., Baschera, G.M., Kohn, J., Kucian, K., Richtmann, V., Grond, U., Gross, M., von Aster, M.: Design and evaluation of the computer-based training program Calcularis for enhancing numerical cognition. Front. Psychol. **4**, 1–13 (2013)
7. Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönmann, C., Plangger, F., Gälli, M., Martin, E., von Aster, M.: Mental number line training in children with developmental dyscalculia. Neuroimage **57**, 782–795 (2011)

8. Siegler, R.S., Fazio, L.K., Bailey, D.H., Zhou, X.: Fractions: The new frontier for theories of numerical development. Trends Cogn. Sci. **17**, 13–19 (2013)
9. National Mathematics Advisory Panel: Foundations for success: The final report of the National Mathematics Advisory Panel (2008)
10. Bailey, D.H., Hoard, M.K., Nugent, L., Geary, D.C.: Competence with fractions predicts gains in mathematics achievement. J. Exp. Child Psychol. **113**, 447–455 (2012)
11. Booth, J.L., Newton, K.J.: Fractions: could they really be the gatekeeper's doorman? Contemp. Educ. Psychol. **37**, 247–253 (2012)
12. Fazio, L.K., Bailey, D.H., Thompson, C., Siegler, R.S.: Relations of different types of numerical magnitude representations to each other and to mathematics achievement. J. Exp. Child Psychol. **123**, 53–72 (2014)
13. Fazio, L.K., Kennedy, C.A., Siegler, R.S.: Improving children's knowledge of fraction magnitudes. PLoS ONE **11**, e0165243 (2016)
14. Ninaus, M., Kiili, K., McMullen, J., Moeller, K.: Assessing fraction knowledge by a digital game. Comput. Hum. Behav. **70**, 197–206 (2017)
15. Kiili, K., Moeller, K., Ninaus, M.: Evaluating the effectiveness of a game-based rational number training - in-game metrics as learning indicators. Comput. Educ. (under Review)
16. Schneider, M., Thompson, C.A., Rittle-Johnson, B.: Associations of magnitude comparison and number line estimation with mathematical competence: a comparative review. In: Lemaire, P. (ed.) Cognitive development from a strategy perspective: A festschrift for Robert S. Siegler. Psychology Press, London (2018)
17. de Hevia, M.D., Girelli, L., Cassia, V.M.: Minds without language represent number through space: Origins of the mental number line. Front. Psychol. **3**, 1–4 (2012)
18. Siegler, R.S., Opfer, J.E.: The development of numerical estimation: evidence for multiple representations of numerical quantity. Psychol. Sci. **14**, 237–243 (2003)
19. Schneider, M., Stern, E.: The developmental relations between conceptual and procedural knowledge: a multimethod approach. Dev. Psychol. **46**, 178–192 (2010)
20. Siegler, R.S., Ramani, G.B.: Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. J. Educ. Psychol. **101**, 545–560 (2009)
21. Moyer, R.S., Landauer, T.K.: Time required for Judgements of Numerical Inequality. Nature **215**, 1519–1520 (1967)
22. Schneider, M., Siegler, R.S.: Representations of the magnitudes of fractions. J. Exp. Psychol. Hum. Percept. Perform. **36**, 1227–1238 (2010)
23. Siegler, R.S., Ramani, G.B.: Playing linear numerical board games promotes low-income children's numerical development. Dev. Sci. **11**, 655–661 (2008)
24. Laski, E.V., Siegler, R.S.: Is 27 a big number? correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. Child Dev. **78**, 1723–1743 (2007)
25. Link, T., Nuerk, H.-C., Moeller, K.: On the relation between the mental number line and arithmetic competencies. Q. J. Exp. Psychol. **67**, 1597–1613 (2014)
26. Siegler, R.S., Thompson, C., Schneider, M.: An integrated theory of whole number and fractions development. Cogn. Psychol. **62**, 273–296 (2011)
27. Fischer, U., Moeller, K., Huber, S., Cress, U., Nuerk, H.-C.: Full-body movement in numerical trainings: a pilot study with an interactive whiteboard. Int. J. Serious Games. **2**, 23–35 (2015)
28. Kiili, K., Ketamo, H.: Evaluating cognitive and affective outcomes of a digital game-based math test. IEEE Trans. Learn. Technol. **PP**(99), p. 1 (2017)
29. R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing (2017). https://www.r-project.org/

30. Wei, T., Simko, V.: corrplot: Visualization of a Correlation Matrix. R package version 0.77 (2016)
31. Wickham, H.: Ggplot2: Elegant Graphics for Data Analysis. Springer, New York (2009)
32. Lindstedt, A., Kiili, K.: Evaluating playing experience and adoption of a math learning game. In: Proceedings of the 1st International GamiFIN Conference, pp. 39–46. CEUR Workshop Proceedings (2017)

Building a Game to Build Competencies

Esther Kuindersma[1(✉)], Jelke van der Pal[1], Jaap van den Herik[2], and Aske Plaat[2]

[1] Netherlands Aerospace Centre (NLR), Amsterdam, The Netherlands
Esther.Kuindersma@NLR.nl
[2] Leiden Institute of Advanced Computer Science, Leiden University, Leiden, The Netherlands

Abstract. Positive developments in aviation, such as the increased safety of air travel, advanced automation and more efficient pilot training courses, may negatively influence the pilots' abilities to handle unknown and unexpected situations. Pilots of the older generation often have experience with manually flying multiple types of aircraft and handling all kinds of incidents. In contrast, the future generation of pilots will not be given the chance to gain the experiences in practice. Consequently, they may not have the competencies needed to handle critical situations. This paper describes how we set about designing a serious game for the acquisition of the essential competencies for critical situations. The game offers meaningful events, which match situations that pilots can face someday. The events trigger and reinforce the competencies of situational awareness, workload management, and application of procedures.

Keywords: Serious games · Zero-fidelity simulation · Aviation Competencies · Critical situations · Meaningful events

1 Serious Games for Aviation Training

Since the introduction of air travel at the beginning of the twentieth century, many advances have been made that have greatly improved the safety of air travel [1]. Aircraft have been made robust, and rules and regulations for air travel have been implemented. Pilot training has become regulated and with the invention of the full flight simulator it has gained in quality and efficiency. The rise of the automatic pilot has also had great effects on the safety. Without any argument, these are positive developments. However, the positive changes may also have a negative side.

Pilots of the older generation often have a background with diverse aviation experiences. Many of them have had a career with a military background or have skills in gliding, aerobatics and other aerospace activities before they started working on large, multi-crew aircraft, building their expertise and preparing them for almost any situation. In contrast, younger airline pilots often take a new, different route to becoming an airline pilot. They start their career in high-tech, modern aircraft with advanced automation, limiting their experience with hands-on, manual flying. They receive extensive training, but they have less hours of flight time when they start their

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 14–24, 2017.
https://doi.org/10.1007/978-3-319-71940-5_2

airline career, and in many cases they have less flying experience in general. Research shows that, in critical situations, pilots fall back on their prior experience [2].

Yet, in such situations younger pilots do not have the experience on which they can fall back. Of course, we know that in normal situations, and even most non-normal[1] situations, extensive training has prepared pilots to perform their jobs adequately, guided by a large number of procedures and checklists. Still, during a flight *unexpected* situations can occur; i.e., situations that do not match with what the pilots anticipate [3]. When a situation is unexpected, and on top of that, *unknown* to the pilot at work, it could (momentarily) surprise him[2], causing him to lose control of the situation which could then turn into an emergency. We refer to such situations as '*critical situations*'. The actual experience problem may manifest itself primarily in these rare situations.

Obviously, the technical knowledge and skills of a pilot are vital, but in critical situations it may happen that the non-technical competencies are key to handling the problem. The pilot's ability to stay calm, think and act is vital to a successful outcome in an emergency. If the required competencies cannot be gained from actual flight experience, they should be developed through training.

Currently, aviation training uses several training methods, such as (1) computer based training, (2) classroom lectures, and (3) flight simulations [4]. Computer-based training (CBT) was introduced into aviation in the 1980's [5] and is now often used in the theoretical part of the initial pilot training and the recurrent training of licensed pilots, together with lectures. Technical skills are commonly trained in flight simulators. And for the real thing, aspirant pilots usually start flying with a single engine aircraft and gradually work their way up to a large airline jet.

Usually, a certified full flight simulator (FFS) is used in training. An FFS is an identical copy of a particular aircraft cockpit, in both physical shape and behaviour. However, the use of an FFS is costly. The time available should be used for training the principal (technical) skills of the pilots. Cognitive skills and competencies are better trained outside these simulators to save money, and possibly to reach a greater effect. This is not to say that simulations cannot be used for the purpose of training cognitive skills, because there are other types of simulations.

In addition to the more traditional training methods, modern training concepts can be considered to supplement the theoretical and simulation-based training, particularly for the training of non-technical skills. One of these concepts is the use of games as a training tool. This is known as game-based learning or '*serious gaming*'.

Games can be used as a training tool for many kinds of learning goals, such as gaining knowledge or developing competencies [6]. The learning goal affects the design of the game. In a serious game, learning usually takes place implicitly [7]. The learner does not receive direct instruction and may not be aware he is learning while playing the game. The learning is an incidental consequence of playing the game. This '*stealth learning*' [8, 9] is in contrast with traditional training methods in which the transfer of knowledge is explicit and the learning goals are clear in advance.

[1] Although 'abnormal' is the familiar opposite of 'normal', 'non-normal' is the term adopted by aviation regulation bodies such as the FAA, EASA and ICAO.

[2] For brevity, we use 'he' and 'him' wherever 'he or she' and 'him or her' are meant.

The learning effect of a serious game is generally attributed to an increase of intrinsic motivation in learners, leading to more attention and longer Time-on-Task [10, 11]. From other publications, we know that the increased motivation is possibly a result of learning in a meaningful context and learning by doing [12]. Improved learning can also be the effect of social aspects, collaboration and competition [13] and of improved retention and transfer [14]. Combining gameplay with a debriefing session, to make the learning explicit by reflecting on it, adds to the learning effect of the game [15].

To build the competencies needed in critical situations, serious games could provide a meaningful alternative to actual flight experience using a so-called '*zero-fidelity simulation*', i.e., a type of simulation in which the target environment is abstracted and the focus is on the cognitive, human aspects of the tasks [16].

This paper starts with a description of the planned study to measure the transfer of the competencies of situational awareness, workload management and application of procedures from a serious game setting to a pilot's performance in an aircraft (Sect. 2). Then it reports on the design of this serious game that implements meaningful events to acquire the essential competencies for critical situations (Sect. 3). The beta version of the Shuttle to Mars game has been submitted to a playtest (Sect. 4). Finally, we discuss the study (Sect. 5).

2 Measuring the Learning Effect of a Serious Game for Pilots

To measure the extent to which a serious game improves the handling of critical situations, we plan to perform an independent-measures experiment (Fig. 1) with two groups of relatively unexperienced airline pilots, to gather both quantitative and qualitative data. We aim to recruit at least 40 participants, men and women, who will be randomly assigned to one of the two groups. Participants in the test group individually play the serious game, described in Sect. 3. The gameplay will be divided into segments that each contains a number of game levels and is concluded with a questionnaire. The questionnaires will contain questions for the assessment of the player's motivation and engagement, as well as for the purpose of debriefing and reflecting on the learning process [15]. The control group will be briefed about the purpose of the study and the importance of the essential competencies during critical situations, and they will fill out a questionnaire. They do not play the game.

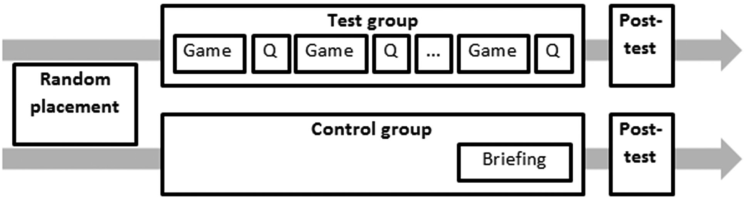


Fig. 1. Serious game transfer experiment procedure

The post-test for each participant will consist of a session in a fixed-base flight simulator for a Boeing 737 aircraft. The participant's performance will be assessed by the examiner focussing on behavioural indicators for the essential competencies.

We expect to see an improvement in the participant's performance during the game (test group) and a difference between the performances of both groups (test and control) in the simulator. This would indicate a positive, transferable learning effect of the serious game. To examine this effect, we will analyse the data coming from (1) the game, (2) the questionnaires, and (3) the assessments made by the examiner in the simulator.

3 Creating Shuttle to Mars

Below, we describe the prototype of a serious game by which the essential competencies for critical situations can be trained. We call this game 'Shuttle to Mars'. In this title, 'Shuttle' refers to the familiarity of a common activity, while 'Mars' appeals to unknown situations and a sense of adventure.

First, the essential competencies that the game will address are identified in subsect. 3.1. Subsequently, we design meaningful events of increasing complexity to create an engaging gameplay and to raise the pilots' level of experience in subsect. 3.2.

3.1 Identifying the Essential Competencies

Airline companies provide their pilots with Standard Operating Procedures (SOP) for a wide range of tasks in normal situations. SOP manuals include many checklists that must be followed in non-normal situations, in which it is not possible to operate the aircraft using the normal procedures.

A non-normal situation is not necessarily an emergency. It may become an emergency when the safety of the aircraft or the persons on board or on the ground is endangered. Non-normal situations happen every day, but rarely result in accidents [17]. Even severe situations do not need to become emergencies, if the pilot acts adequately. The main task of the pilot is to keep control over the aircraft at all times. A safe continuation of the flight must be secured before accomplishing any non-normal checklist or attempting to handle and solve the problem. A pilot needs certain competencies in order to be able to act accordingly.

The International Civil Aviation Organization (ICAO) has provided a list (Table 1) of eight core competencies describing the technical and non-technical knowledge, skills and attitudes that are needed to operate safely, effectively and efficiently in a commercial air transport environment [18].

Table 1. ICAO Core Competencies

1. Application of Procedures (AP)
2. Communication (COM)
3. Aircraft Flight Path Management, automation (FPM)
4. Aircraft Flight Path Management, manual control (FPM)
5. Leadership and Teamwork (L&T)
6. Problem Solving and Decision Making (PS&DM)
7. Situation Awareness (SA)
8. Workload Management (WM)

To determine which of the ICAO competencies are essential in critical situations, we performed (1) a document study, (2) an observation, and (3) a series of interviews. The document study served to find both formal and informal sources on the tasks that a pilot has to perform during normal and non-normal situations. This study yielded scientific articles, public documents, websites and weblogs. In these materials multiple sets of guidelines for handling non-normal situations [19–21] were found. Although the phrasing of the guidelines varies, eleven common guidelines have been identified and matched to the ICAO core competencies (Table 2).

Table 2. Matching common handling guidelines with ICAO competencies

Handling guideline	ICAO competencies
Be aware of changes in situation	7. Situation Awareness
Perform your primary task	1. Application of procedures 3. Flight Path Management, automation 4. Flight Path Management, manual control
Stay calm	8. Workload Management
Identify source of the problem	6. Problem Solving and Decision Making
Determine severity	6. Problem Solving and Decision Making
Come up with a plan	6. Problem Solving and Decision Making
Prioritize duties	6. Problem Solving and Decision Making 8. Workload Management
Delegate duties	8. Workload Management
Use non-normal checklists	6. Problem Solving and Decision Making
Take action	6. Problem Solving and Decision Making
Communicate	2. Communication

A simulator training session in a Boeing 777 FFS was informally observed. The training session consisted of a large number of non-normal situations and emergencies initiated by the instructor/evaluator. The two pilots were scored on communications, performance of technical operations, and their application of procedures and checklists. The way the pilots handled the situations in the simulator corresponded with the handling guidelines (Table 2) found in the document study.

Semi-structured interviews were conducted with four captains and one first officer from several airlines; 4 men and 1 woman. Each interview lasted about 60 min. Their work experience ranged from 15 to 25 years and from 5,000 to 14,000 flight hours. The questions focused on the pilots' backgrounds, their experience with non-normal situations and their view on the essential competencies, based on the ICAO core competencies (Table 1). The interviews confirmed that real emergencies are rare. All pilots had only experienced a few non-normal situations in their careers and they did not consider most of these situations an emergency. They all indicated that flying the plane (competencies 3 and 4) is most important, together with Problem Solving and Decision Making (competency 6), Situational Awareness (competency 7) and Workload Management (competency 8). This is also consistent with the guidelines found in the document study, as shown in Table 2.

Although the study included a small number of interviews and one observation, we believe the findings are valid, based on the methodological triangulation. Therefore, we established here that the following four competencies are essential in critical situations:

1. Situation Awareness
2. Workload Management
3. Application of Procedures
4. Problem Solving and Decision Making

The Shuttle to Mars game will aim to activate these competencies, in order to improve the pilots' ability to act adequately in critical situations.

3.2 Designing the Game

Shuttle to Mars is designed to build the essential competencies needed by airline pilots in critical situations without explicitly telling the players about it [8]. It aims to create a positive attitude toward these competencies and lower the threshold of applying them in their daily jobs. The game is developed as a single-player first-person cockpit adventure for iPad with the Unity 3D platform. It should engage the players and appeal to them to play the game out of their own accord.

The narrative of Shuttle to Mars is set in outer space. This space theme was chosen for the similarities between space travel and air travel, which are distant but still easily recognized by the player. Although the game does not offer a recognizable airplane cockpit and the tasks do not resemble actual piloting, it is in fact a '*zero-fidelity*' simulation of flying a commercial airplane [15]. The individual aspects of tasks and situations in the game correspond with those of flying. Game mechanics are combined into '*meaningful events*'; i.e., complex situations that have a strong link to critical situations in aviation. While composing meaningful events from the game mechanics available, emphasis lies on triggering and reinforcing the essential competencies.

The space theme also allows for a motivating storyline, a continuous primary task, a high workload with secondary tasks, and opportunities for surprising situations.

In the game, the player takes on the role of the captain of a Mars shuttle. His job is to transport cargo through outer space. Each delivery run is presented as a mission (level). The player's goal in each mission is to reach the destination, with as little damage as possible to cargo, crew and spaceship. In order to reach its destination with the highest score possible, the spaceship has to be controlled, resources need to be managed, safe passage through all space sectors has to be arranged and all kinds of situations need to be dealt with. The player needs to stay calm, stay focused and use problem solving skills to succeed.

The complexity of the game increases in the course of the missions. In total the game will comprise twenty missions. The first three missions aim to familiarize the player with the game and the possibilities of the gameplay, and to let him develop some routine in performing the tasks. The rest of the missions serve to help the player reach the learning objective.

Fig. 2. Shuttle to Mars cockpit

The game environment is dynamic and requires much interaction from the player. It contains elements of surprise that aim to throw the player off balance. From his position in the cockpit, the player has a first person view of the dashboard, the overhead panel, the Head Up Display (HUD) and, through the windshield, outer space (Fig. 2). As part of the storyline the player will interact with non-player characters (NPC); two crew members, Galaxy Traffic Control, other spaceships and potential enemies. The player will receive messages, requests and orders through onscreen and audio notifications. He can respond by giving input through a numeric keypad and through buttons, switches and sliders.

The player's primary task is to stay on track and avoid obstacles, along with monitoring the spaceship's status. During each mission meaningful events arise (Table 3) that the player has to respond to and that call upon the essential competencies to be able to perform certain tasks. Notifications and signals must be dealt with in a timely manner to prevent the situation from deteriorating into a catastrophe. Situations range from relatively simple to very complex and difficult to handle.

Table 3. Examples of meaningful events, real life parallels and competencies

Meaningful event	Parallel with real life	Competency
Follow designated route	Routinely perform primary task combined with secondary tasks	FPM
Interactions with NPC	Communicate with several sources, through several media	COM
Authentication call	Perform procedure	AP, WM
System failures	Maintain level of situation awareness, Delegate work	SA, WM
Asteroids on path	Choose best solution	PS&DM
Resource shortages	Maintain level of situation awareness, Balance resources	SA, PS&DM
Pirate encounter	Weigh options	PS&DM
Passing space ships	Maintain level of situation awareness	SA

In critical situations in aviation, complicated incidents are combined with dangerous circumstances and time pressure. The meaningful events in the game mimic this. They are matched, based on four elements: (1) the essential competencies, (2) their behavioural indicators, (3) working conditions and (4) characteristics of the tasks. The parallels created are intended to stimulate transfer of the competencies from the game to the actual work environment. To ascertain relevant parallels, airline instructors are consulted during the design of the meaningful events.

4 Playtesting the Game

The beta version of Shuttle to Mars was submitted to a playtest. The purpose of a playtest is to determine whether the game produces the experience which the designers intended to reach [22]. A playtest also identifies pacing and balancing problems [23]. In usability studies, a sample of five users will identify almost as many problems as a larger group [24]. For a serious game, playtesting is done to determine whether the target audience will be engaged in the game. Without engagement, the educational objective will not be reached. Furthermore, playtesting provides information as to whether the game controls are understandable and tractable, and the game is playable for the target group. The game's functionality and technical quality are not tested in a playtest [22].

Five male participants played Shuttle to Mars; four airline pilots and one flight simulator engineer. They played selected parts of the game, and were asked to think aloud during the entire playtest. The playtest was directed by a supervisor and observed by two observers. Video and audio recording were made and the iPad screen was captured. During the playtest the supervisor asked questions about the player's behaviour in the game, when the participant had completed a mission.

Before and after playing the game, the participants were interviewed and they filled out a questionnaire with 21 questions using a five-point scale. Nine questions were aggregated into subgroups regarding four important topics: (1) controls, (2) enjoyment, (3) engagement, and (4) parallels with reality. The results (Fig. 3) show that the game controls were understandable and tractable. Moreover, the participants enjoyed the game. They became engaged in the game and tried to succeed. All participants acknowledged the parallels between the space-themed storyline and aviation reality. The outcomes correspond with the participants' commentaries during the test. On top of that, three of the four airline pilots spontaneously identified the competencies which the game aims to reinforce.

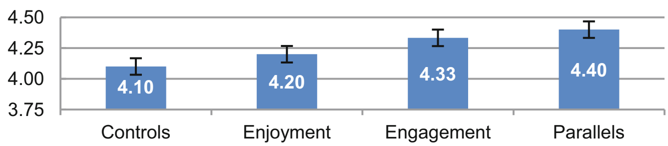


Fig. 3. Questionnaire results (five point Likert scale, means with standard errors)

The playtest was aimed at establishing the playability of Shuttle to Mars, not at finding statistical proof of its learning effect. The outcomes of the playtest give us confidence that Shuttle to Mars will be fun to play and will possibly have a positive learning effect. More missions will be designed to provide learning through meaningful events. With the completed game we will have a second playtest with another five participants. That playtest will focus on the quality of the game data, but will also provide insight into the appreciation of the gameplay. Afterwards, we will move forward to examine the learning effect and the learning transfer in the experimental setup (see Sect. 2).

5 Discussion

The goal of professional training is to improve performance in the actual work setting. The behaviour, knowledge and skills acquired in a specific training environment must be transferred to situations in a different, real world environment [25]. Transfer of learning from a training situation to a working context is hard to prove. In case of handling critical situations in aviation, proving it is even impossible due to safety and ethics. Obviously, sending pilots up in the air and creating a critical situation to see how they fare is ethically unacceptable. However, training in a flight simulator is safe, and the transfer from flight simulator to cockpit has long been acknowledged [26]. Therefore, we resort to an experiment involving a flight simulator. We will look for evidence that the Shuttle to Mars game has a positive effect on the adequate handling of a critical situation in the flight simulator, as an indication that it will also have a positive effect on the pilot's performance in the actual aircraft.

The game data will provide information about the player's performance, including response times, correct answers, and procedures followed. We expect that the participants will get better at handling the meaningful events as they spend more time playing the game. The game data is connected to behaviour indicative of the essential competencies. The questionnaires will provide information about (1) the participant's motivation during the gameplay, (2) his engagement in the game, and (3) his understanding of the learning objectives and the relevance of the game. As an active, positive attitude is beneficial for the learning effect [27], the questionnaires continuously monitor the participants' engagement and motivation levels.

We expect to see a stronger improvement in simulator performance by the game playing participants in comparison to the control group. Such a difference may indicate that playing the game improves the player's ability to handle critical situations in the simulator, and consequently improve performance in the work environment.

The Shuttle to Mars prototype has the potential of becoming an effective training (research) tool. Before the experiment starts, some improvements need to be implemented. The game levels need to be filled with meaningful events that allow the building of the essential competencies. When a final playtest is successful, the experiment can be performed to examine the learning effect and to what extent this learning transfers from the game to the work environment.

References

1. Collins, S.: Safer Skies. Allianz Global Corporate & Specialty, vol. 1, pp. 22–24 (2015)
2. Stepniczka, I., Tomova L., Rankin, A., Woltjer, R., Sladky, R., Tik, M.: D3.1 Final Analysis of Research Evaluation. Man4Gen consortium, Vienna (2015)
3. Rankin, A., Woltjer, R., Field, J., Woods, D.: "Staying ahead of the aircraft" and Managing Surprise in Modern Airliners. In: Paper presented at 5th Resilience Engineering Symposium: Managing trade-offs, Soesterberg, The Netherlands (2013)
4. European Cockpit Association: Pilot Training Compass: Back to the future (2013)
5. Franks, P., Hay, S., Mavin, T.: Can competency-based training fly? an overview of key issues for "Ab Initio" Pilot Training. Int. J. Training Res. **12**(2), 132–147 (2014)
6. Wouters, P., Van der Spek, E.D., Van Oostendorp, H.: Current practices in serious game research: A review from a learning outcomes perspective. In: Connolly, T.M., Stansfield, M., Boyle, E. (eds.) Games-based learning advancements for multi-sensory human computer interfaces: techniques and effective practices, pp. 232–250. IGI Global (2009)
7. Gee, J.P.: Deep learning properties of good digital games: how far can they go. In: Ritterfeld, U., Cody, M., Vorderer, P.A. (eds.) Serious Games: Mechanisms and Effects, pp. 67–80. Routledge, New York (2009)
8. Sharp, L.A.: Stealth learning: unexpected learning opportunities through games. J. Instruct. Res. **1**, 42–48 (2012)
9. Annetta, L.A.: Video games in education: Why they should be used and how they are being used. Theor. Pract. **47**(3), 229–239 (2008)
10. Mautone, T., Spiker, V.A., Karp, M.R., Conkey, C.: Using games to accelerate aircrew cognitive training. In: Paper presented at Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL (2010)
11. Berliner, D.C.: What's all the fuss about instructional time? In: Ben-Peretz, M., Bromme, R. (eds.) The Nature of Time in School. Theoretical Concepts, Practitioners Perceptions, pp. 3–35. Columbia University Teachers College Press, New York (1990)
12. Garris, R., Ahlers, R., Driskell, J.: Games, motivation, and learning: a research and practice model. Simul. Gaming **33**(4), 441–467 (2002)
13. Gee, J.P.: Learning and games. In: Salen, K. (ed.) The Ecology of Games: Connecting Youth, Games, and Learning, pp. 21–40. MIT Press, Cambridge (2008)
14. Knulst, M.: Serious Gaming & Didactics: a Review on Game, Instructional, and Player Variables in Serious Game Design. NLR-TR-2014-397. Netherlands Aerospace Centre, Amsterdam (2014)
15. Crookall, D.: Serious Games, Debriefing, and Simulation/Gaming as a Discipline. Simulation & Gaming **41**(6), 898–920 (2010)
16. Toups, Z.O., Kerne, A., Hamilton, W.A.: The team coordination game: zero-fidelity simulation abstracted from fire emergency response practice. ACM Trans. Comput. Hum. Interact. **18**(4), 23 (2011)
17. Burian, B.K., Barshi, I., Dismukes, R.K.: The Challenges of Aviation Emergency and Abnormal Situations. NASA Technical Memorandum 2005-213462. NASA Ames Research Center, Moffett Field, CA (2005)
18. International Civil Aviation Organization: Manual of Evidence-based Training. Doc 9995. ICAO, Montreal (2013). http://skybrary.aero/bookshelf/books/3177.pdf
19. Owens, B.: Handling an emergency (2012). http://iflyblog.com/handling-an-emergency
20. Hainan Airlines: Standard Flight Operations Manual SOP 737-800. Hainan Airlines Company Limited, Haikou (2009)

21. Kahn, K.M.: Emergency Exit: How to Handle Non-Normal Events (2004). http://www.aopa.org/news-and-media/all-news/2004/december/flight-training-magazine/emergency-exit
22. Becker, K., Parker, J.: Methods of design: an overview of game design techniques. In: Schrier, K. (ed.) Learning, Education and Games: Volume One: Curricular and Design Considerations, pp. 179–198. ETC Press, Pittsburgh (2014)
23. Desurvire, H., El-Nasr, M.S.: Methods for game user research: studying player behavior to enhance game design. IEEE Comput. Graph. Appl. **33**(4), 82–87 (2013)
24. Nielsen, J.: Why you only need to test with 5 users (2000). https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users
25. Yamnill, S., McLean, G.N.: Theories supporting transfer of training. Hum. Resour. Dev. Q. **12**(2), 195–208 (2001)
26. Allerton, D.: Principles of Flight Simulation. Wiley, Chichester (2009)
27. Wouters, P., Van Nimwegen, C., Van Oostendorp, H., Van Der Spek, E.: Meta-Analysis of the Cognitive and Motivational Effects of Serious Games. J. Educ. Psychol. **105**(2), 249–265 (2013)

Training of Spatial Abilities with Digital Games: Impact on Mathematics Performance of Primary School Students

Laura Freina(✉), Rosa Bottino, Lucia Ferlino, and Mauro Tavella

National Research Council, Institute for Educational Technologies,
via de Marini, 6, 16149 Genova, Italy
{freina, bottino, ferlino, tavella}@itd.cnr.it

Abstract. The “A me gli occhi project” has been organized to support the development and consolidation of visuospatial abilities in students of the last two years of primary school, with the hypothesis that such a training would have a positive impact on their performance in mathematics. The present paper focuses on the second phase of the project, during which twenty game based training sessions have been organized along the whole school year. Students’ mathematical abilities have been measured at the beginning of the project and at the end using a standardized math test. Results show some improvements in the experimental group compared to the control one, confirming the initial hypothesis.

1 Introduction

Several studies have demonstrated that there is a strong correlation between visuospatial abilities and school performances in Science, Technology, Engineering, and Mathematics (STEM) [1]; people who have good visuospatial abilities are also more likely to engage in STEM studies and later in STEM related jobs [2–4]. There are great differences among people in visuospatial abilities; nevertheless, these abilities can be trained with long lasting improvements affecting academic success in STEM subjects [5]. Moreover, there are evidences [6], that early interventions focused on the development and training of visuospatial abilities are particularly useful since such abilities develop slowly and primary school usually offers a more flexible context, both in terms of curricula constraints and time organization, to introduce non-traditional activities. Unfortunately, there appears not to be enough attention to the development of these abilities in formal education.

This paper describes “A me gli occhi” (“Eyes on Me”), a two-year project aimed at improving primary students’ performance in mathematics through training visuospatial abilities by means of games, mainly digital ones. This project stems from previous in-field research studies carried out by a group of researchers of ITD (Institute of Educational Technology of the Italian National Research Council) with the aim of investigating the role digital games can have in stimulating and assessing primary school students’ logical and reasoning skills [7, 8].

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 25–40, 2017.
https://doi.org/10.1007/978-3-319-71940-5_3

In the following, a brief introduction of what is usually meant with “visuospatial abilities” is provided, and then the “A me gli occhi” project and its main outcomes are described. In the conclusions, some ideas for future research are outlined.

2 Visuospatial Abilities and Digital Games

As reported by Bednarz and Lee [9], there is not a unique and globally accepted definition of visuospatial abilities; nevertheless, most researchers agree to include in such a definition spatial visualization (the ability to mentally represent and operate on visual stimuli), spatial orientation (the ability to picture spatially arrayed elements from different perspectives), and the understanding of spatial relations (spatial distributions, locations, imagine maps from verbal descriptions and in general work with maps). Furthermore, visual working memory also plays a key role in visuospatial abilities [10].

There is an extensive body of research that has accumulated over the past twenty years, which consistently shows a correlation between visuospatial abilities and success in mathematics and science. For example, Newcombe [6] reports several different longitudinal studies that started back in the fifties and followed the development of a large number of American children from nursery school to adulthood. These studies have shown that there is a strong correlation between visuospatial abilities and results in STEM areas. Moreover, they indicated that having good visuospatial abilities increases the probability to undertake STEM related jobs.

Research has shown that visuospatial abilities can be improved with practice and that such an improvement has long-term effects on student learning [3]. At university level, it has been shown that visuospatial training significantly “boosts” achievements in STEM disciplines by helping students persist in early challenging course work [4]. At an earlier age, enhancing children visuospatial abilities can contribute to improve their mathematical skills [6] and, consequently, to prevent their future failure in this discipline.

The ability to manipulate quickly and efficiently the spatial properties of objects and the spatial relations among them is a key skill in today’s world, which is increasingly demanding from a visual point of view, relying more and more on core graphic representations rather than text-based representations. For example, as Mulligan [11] notices, interfaces are becoming less and less alphanumeric and more visuospatial, requiring people to be proficient in spatial understanding.

Some research studies support the assumption that visuospatial abilities training can be effective using a games-based approach [12]. Actually, the use of games in education can enhance students’ motivation and involvement in the activities and can help in keeping their attention high [13, 14]. This causes the students to put a greater effort in the given tasks and, therefore, increase significantly the foreseen advantages. Moreover, experiences that we have conducted since many years have highlighted the high potential of digital games to develop and strengthen reasoning and visual skills and have shown the positive impact of their use on school performance [15].

3 The "A me gli occhi" Project

Even if it is widely accepted that visuospatial abilities have a direct impact of school performance, there appears not to be enough attention to the development of these abilities in formal education [16]. The project has therefore been set up in order to verify if a direct intervention at an early school age would actually have a positive influence on mathematics outcomes.

A specific training has been organized with the aim of improving and consolidating visuospatial abilities, as well as exercising the visuospatial working memory in children in the last years of primary school. The hypothesis underlying the project has been that training visuospatial abilities of young students would have a positive impact on their school performance in math. The approach that has been selected for the training sessions is based mostly on the use of digital games.

Surtees et al. [17] demonstrated that visuospatial abilities stem from embodied processes: those parts of the brain that are involved with movements of the body in space are activated when performing spatial reasoning tasks. Due to this characteristic, training in an immersive virtual world may have a positive impact on performance. Virtual worlds can offer different levels of "immersion", by surrounding the users making them feel as if they were "really there".

The project has so been divided into two phases: the first aimed at identifying the best immersion level needed to maximise the players' performance in tasks involving visuospatial abilities, the second oriented to measure the impact of an extended game based training on school results in math.

During the first phase of the project, a virtual reality game, "In Your Eyes", specifically implemented to support the development and consolidation of some visuospatial abilities has been used with a group of about 100 students to assess the impact of the different possible levels of immersion on the performance in the task. The game was available in three different versions, each corresponding to a different level of immersion in a virtual world. Results show that, while having the possibility to move freely in a virtual environment helps to have a better performance in the task, the complete immersion using a Head Mounted Display does not have a significant impact [18]. For this reason, in the second phase of the project the complete immersion in a virtual world has not been used.

The present paper is focused on the second part of the project. Its experimental design, the target population, the pre-text and post-test used, the training activities and the data analysis performed are described together with a discussion of the results obtained.

3.1 Aim of the Project

The project aim has been to verify if a specific training of visuospatial abilities could have a positive impact on primary school students' mathematics performance. In our hypothesis, visuospatial abilities should get better with training and, therefore, school results in mathematics should improve.

3.2 Target Population

The age of the target population has been accurately selected. According to Meyer's experiment [19], starting from around the age of 8, the visuospatial representations in the working memory play an increasingly important role in predicting mathematical reasoning and numerical operations achievements, while in younger children these abilities are more related to language proficiencies. Furthermore, several studies seem to indicate that visuospatial abilities complete their development around the age of 8 [17, 20, 21]. It has therefore been decided to address primary students from grades 4 and 5 (aged 9–11) in the project.

Four classes from the "Scuola Primaria Cantore", a local primary school in Genova (Italy) took part in the second phase of the project. All the students of each class have been involved. The experimental group included 38 students, 23 of which from a grade 4 class and 15 from a grade 5 class. The control group included 41 students, 25 from a grade 4 class and 16 from a grade 5 class. All the students were between age 9 and 11. Even if different classes had different teachers, they all followed the same program and regular meetings were organized along the school year within the school to guarantee teaching homogeneity.

All training sessions took place at the school premises during school hours as part of the normal activities. Parents and students have been informed before the start of the project.

3.3 Research Method

A quasi-experimental research has been set up with the experimental and the control group of participants. It has not been possible to randomly assign participants to the two groups since classes could not be divided. A pre-test and a post-test measuring the students' performance in math have been collected from both groups. Twenty game based training sessions have been organized with the experimental group in the period between November 2016 and May 2017 (one session a week), while the control group just followed the traditional school activities.

3.4 Pre and Post Test

The "AC-MT 6-11", a standardized math test for the Italian primary school has been selected for the project [22]. The test is standardized for each primary grade both at the beginning of the school year and at the end. There are different versions of the test for each school grade, all organized into five different sections. Scores for each section have been kept separated in the data analysis because the test does not give an overall evaluation, each part of the test involves different abilities and scores are different and cannot be simply added.

All the participating classes have taken the tests both in November (pre-test) and in May (post-test). Class teachers managed them as a normal school tests. Scores from the pre-test and the post-test are discussed in the "Data analysis" chapter.

In the following the different sections, of the math test used in the project are briefly described.

- **Written number operations**. This section was made of eight operations, including sums, differences, multiplications and divisions. Each operation was worth one point, for a maximum score of 8.
- **Identifying the biggest between two numbers with decimals**. The section involved the comparison between six pairs of numbers for a maximum score of 6.
- **Deducing a number from word names using the place value rules**. This section included six exercises, for a maximum score of 6.
- **Ordering numbers with decimals**. This task included 10 exercises. The first five exercises asked to order four numbers from the smallest to the biggest, while the following five asked the other way round. The maximum possible score was 10.
- **Word problems**. The session on word problems included five problems, four of which needed a twostep reasoning process to be solved, while the other one was simpler and needed only one. The overall maximum score was 9.

3.5 Training of the Visuospatial Abilities

The training sessions have been organized once a week, always with the presence of a researcher from ITD-CNR and a prospective teacher (two senior student teachers have been included in the project, they were present during the training sessions as part of their apprenticeship). During the training, the experimenter worked with half of the class (groups of 8–12 students) for 45 min per group, while the rest of the students continued class activities with the teacher, at the end of the period the groups swapped. Each student worked individually with an Android tablet. The training included several different activities: playing with digital games, a Minecraft contest and some gaming activities not based on digital games.

Digital games. Several game apps and computer games have been accurately chosen according to their characteristics in order to focus on the training of visuospatial abilities (see Annex A for a list of the games used during the project). The activities proposed to students were mainly of the following types:

- Visual memory activities, which included activities like remembering a list of objects, an ordered sequence of elements, the location of items on a 2D board, etc.
- Movement and rotation of 2D objects by playing traditional games like jigsaw puzzles, tangram, etc.
- Definition of a path on a 2D plane by playing several maze based games.
- Axial symmetry in games in which an image or a solid had to be reflected as on a mirror.
- Movement and rotation of 3D objects using games in which the player had to follow a path in a 3D structure. The interface of the games allowed to turn around the structure and to analyse it from different perspectives.
- Spatial Perspective Taking (SPT: the ability to predict what a scene would look like from a different perspective) using the game “In Your Eyes” [23].

Minecraft contest. Students have been given the plan of a two-floor house with one bedroom, a living room, a bookshelf and a balcony. The plan was manually sketched on paper. Students were free to build the house using the Android version of Minecraft according to the instructions, adding any further element to their wish, as long as they kept the basic items. The works have then been selected: only those whose houses had the characteristics defined by the original project have been admitted to the finals. The finalists have then been voted by the students of the other experimental class.

Concrete gaming activities not based on digital games. Three activities were not strictly based on digital games:

- **SPT game** – a game addressing the SPT ability has been set up in a classroom using square desks as shown in Fig. 1. The experimenter sits on the chair while the player, standing on a white placeholder on the floor, is given four pictures of the table seen from the four sides. S/he has to pick the one that represents the scene from the experimenter's point of view.

Fig. 1. The classroom version of the "In Your Eyes" game.

- Origami – students followed instructions on their tablet while building the figures out of real paper. The origami activity develops visuospatial abilities in that it forces the student to understand how to fold a flat surface into a 3D object, building, while doing so, a mental understanding of the object.
- Room plans – students learned to recognize the plan of a terrace and the location of some items on it, and then they drew the plan of their own classroom. Dealing with plans means understanding a 2D projection of a 3D space.
- Photo activity – each student was given a picture of the room where s/he was and s/he had to move to the place where the picture had been taken from. This task requires the participants to be able to imagine what the room they are in looks like when seen from another point of view [6].

3.6 Data Analysis

Data from the math pre and post-test have been statistically analysed. Since the test used was divided into five different sections, which were rated differently, they have been analysed separately.

For each part of the math test, a complete factorial analysis with a mixed linear model has been applied. Two within-subjects independent variables have been considered: group (experimental vs control) and grade (4 vs 5), and a within-subjects variable: the time when the math test has been done (pre-test vs post-test).

Written Number Operations. The descriptive statistics are reported in Table 1.

Table 1. Descriptive statistics (average ± standard deviation) of the score for the written operations test related to the group (experimental vs. control), and grade. Max score 8.

Grade	Group	Time	Written operations
Four	Control (n = 25)	Pre-test	6.10 ± 1.80
		Post-test	6.83 ± 1.12
	Exp. (n = 23)	Pre-test	6.33 ± 1.83
		Post-test	6.70 ± 1.39
Five	Control (n = 16)	Pre-test	6.56 ± 1.08
		Post-test	5.77 ± 2.10
	Exp. (n = 15)	Pre-test	5.20 ± 2.56
		Post-test	6.47 ± 1.72

For the Written Number Operations, a statistically significant group by class by time interaction has been found ($F(1, 70.77) = 8.04$, $p = .006$) (Fig. 2). Post-hoc tests on the estimated marginal means show that the effect is due to:

- higher scores in grade 5 control group compared to the experimental group at the pre-test ($t(125.64) = 2.03$, $p = .045$);
- higher scores at the post-tests compared to the pre-tests for the experimental grade 5 group ($t(69.24) = 2.68$, $p = .009$);
- higher scores for grade 4 compared to grade 5 in the experimental group at the pre-test ($t(124.80) = 2.00$, $p = .047$).

A closer look at data shows that, in all cases but grade 5 control group, grades were better in May than in November, and the standard deviation decreased. This show a trend towards better scores and a greater homogeneity of the class. In particular, the experimental grade 5 has been able to regain an initial disadvantage measured at the pre-test.

Identifying the biggest between two numbers with decimals. Table 2 shows the descriptive statistics for this section. No statistically significant effect has been found (Fig. 3), nevertheless, trends are very interesting in this section. Actually, in both the experimental classes, the average score at the post-test is higher than in the pre-test,

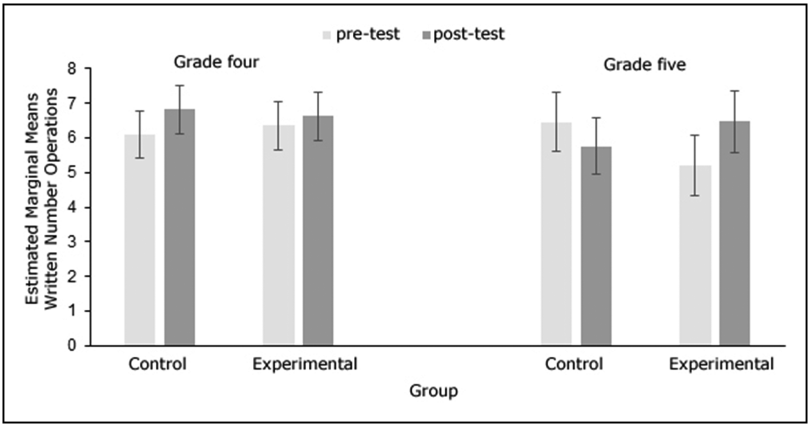


Fig. 2. Estimated Marginal means for the written number operations. Whiskers represent confidence intervals.

Table 2. Descriptive statistics (average ± standard deviation) of the score for the "Identify the Bigger Number" test related to the group (experimental vs. control), and grade. Max score 6.

Grade	Group	Time	Identify the bigger number
Four	Control (n = 25)	Pre-test	5.60 ± 1.26
		Post-test	5.30 ± 1.77
	Exp. (n = 23)	Pre-test	5.57 ± 1.31
		Post-test	5.83 ± 0.39
Five	Control (n = 16)	Pre-test	5.63 ± 1.50
		Post-test	5.59 ± 1.46
	Exp. (n = 15)	Pre-test	5.47 ± 1.55
		Post-test	5.80 ± 0.41

furthermore, the standard deviation decreases a lot, suggesting that the whole class may have taken advantage from the training of their visuospatial abilities.

Deducing a number from words. Table 3 reports the descriptive statistics. In this task, a statistically significant effect between grade and time has been found (F(1, 74.97) = 5.10, p = .027). Post hoc tests on the Estimated Marginal Means have shown that the effect was due to the fact that scores in both grade 4 classes were higher in the post-test than the pre-test (t(75.84) = 2.05, p = .043), and grades in grade 4 classes were higher than those in grade 5 classes (t(75.84) = 2.08, p = .040) (Fig. 4).

Grade 4 classes show a greater improvement in this task when compared to performance of grade 5s. Furthermore, grade 4 post-tests show better scores than the pre-tests. This may also be related to the different grades: learning place value rules is usually dealt with in grades three and four, and this is bound to have a greater impact on students' improvements along the school year.

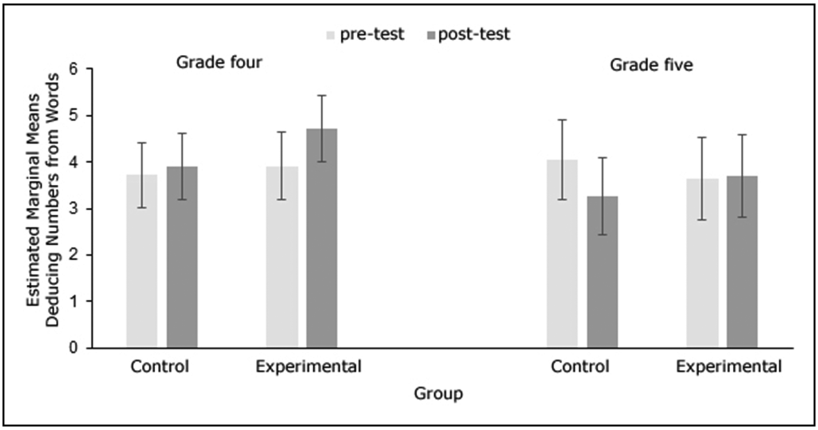


Fig. 3. Estimated Marginal means for the identifying the biggest number task. Whiskers represent confidence intervals.

Table 3. Descriptive statistics (average ± standard deviation) of the score for the deducing a number from words test related to the group (experimental vs. control), and grade. Max score 6.

Grade	Group	Time	Deducing a number from words
Four	Control (n = 25)	Pre-test	3.72 ± 1.54
		Post-test	3.87 ± 1.87
	Exp. (n = 23)	Pre-test	3.94 ± 1.26
		Post-test	4.72 ± 1.19
Five	Control (n = 16)	Pre-test	4.06 ± 1.17
		Post-test	3.27 ± 2.02
	Exp. (n = 15)	Pre-test	3.63 ± 2.33
		Post-test	3.70 ± 2.60

Ordering. Descriptive statistics are reports in Table 4. In the ordering task, the only significant effect was the interaction between group and time ($F(1, 76.08) = 5.69$, $p = .020$). Post hoc tests on the Estimated Marginal Means have shown that for both grades, scores at the post test were higher in the experimental group than in the control group ($t(144.85) = 2.74$, $p = .007$), and the scores of the experimental group were higher at the post-test than at the pre-test ($t(75.47) = 2.64$, $p = .010$) (Fig. 5).

This part of the math test is the one that mostly shows the impact of the training of visuospatial abilities. In fact, this task relies on the visuospatial working memory [24], which has been trained during the project both with specifically aimed digital games, and as a side effect of the general visuospatial abilities.

Word Problems. In the word problems, the main effect of time was significant ($F(1, 74.57) = 12.29$, $p = .001$), due to lower scores in the pre-test than in the post-test. Furthermore, the effect of group and time was significant ($F(1, 74.57) = 11.51$,

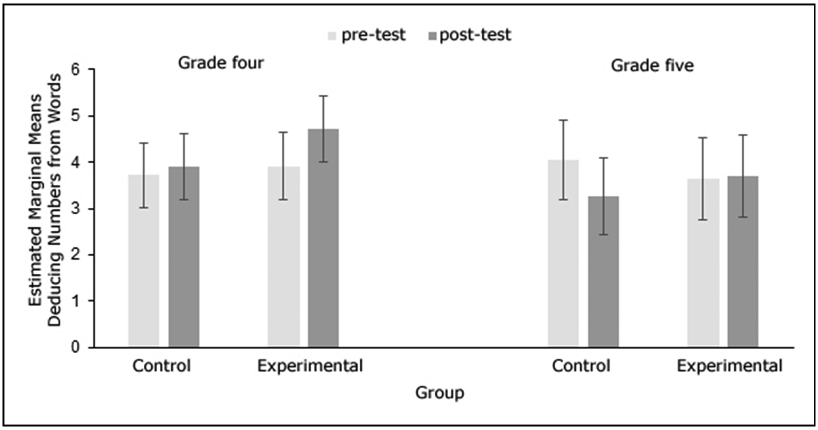


Fig. 4. Estimated Marginal means for the deducing numbers from words. Whiskers represent confidence intervals.

Table 4. Descriptive statistics (average ± standard deviation) of the score for the ordering test related to the group (experimental vs. control), and grade. Max score 10.

Grade	Group	Time	Ordering
Four	Control (n = 25)	Pre-test	7.46 ± 1.54
		Post-test	7.59 ± 2.67
	Exp. (n = 23)	Pre-test	7.41 ± 2.42
		Post-test	8.22 ± 1.17
Five	Control (n = 16)	Pre-test	6.97 ± 1.06
		Post-test	6.29 ± 3.10
	Exp. (n = 15)	Pre-test	6.73 ± 1.78
		Post-test	8.20 ± 1.40

p = .001). Post hoc tests on the Estimated Marginal Means showed that the effect was due to higher scores in the post-test than in the pre-test for the experimental group (t (74.16) = 4.08, p < .001) (Fig. 6).

In this part of the test, a significant impact has been found: scores were higher in the post-test than in the pre-test in the experimental group, while control groups did not change much. This may be due to the positive impact of the increased visuospatial abilities and their role in the interpretation and mental representation of the problem to be solved (Table 5).

3.7 Questionnaire on Game Enjoyment

In order to assess students' enjoyment of the single games, data has been collected using specifically developed questionnaires.

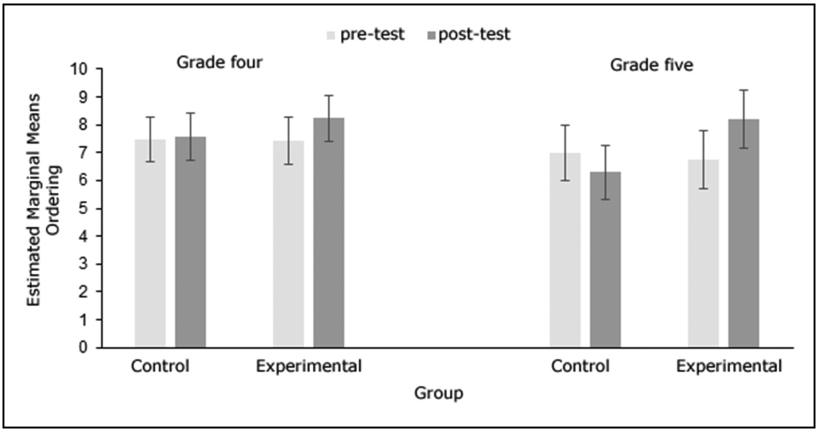


Fig. 5. Estimated Marginal means for the ordering task. Whiskers represent confidence intervals.

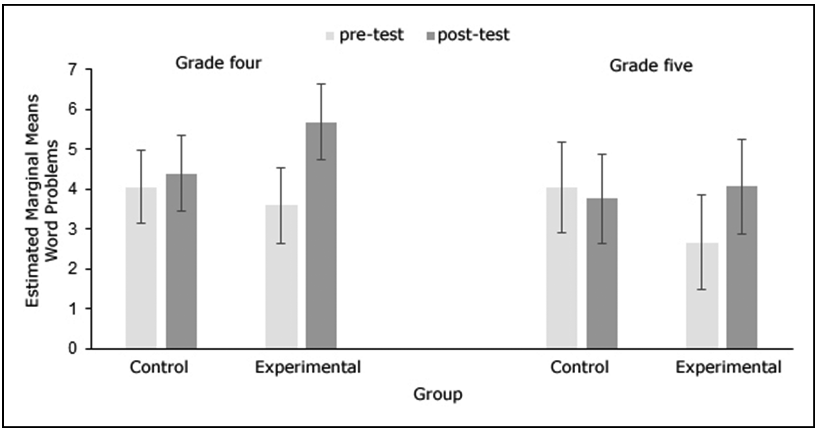


Fig. 6. Estimated Marginal means for word problems. Whiskers represent confidence intervals.

Students have been asked, if they wished, to state freely the preferred games, and 82.05% of participants answered the specific question. It is worth noticing that the Minecraft contest was mentioned by 66.67% of the students, while the other games have been mentioned very little (Mekorama 23.08%, Origami 20.51%, all the others less than 6%).

In the same questionnaire, students were also asked to select the type of activity they enjoyed best between digital games, origami, room plans, and photo activity (more than one choice was possible). Most participants (87.18%) affirmed that they preferred playing with the digital games, only 30.77% preferred Origami, while the other two activities obtained very little interest: 15.38% for the room plans and 7.69% for the photo activity.

Table 5. Descriptive statistics (average ± standard deviation) of the score for the word problem test related to the group (experimental vs. control), and grade. Max score 9.

Grade	Group	Time	Word Problems
Four	Control (n = 25)	Pre-test	4.06 ± 2.32
		Post-test	4.41 ± 2.61
	Exp. (n = 23)	Pre-test	3.61 ± 2.47
		Post-test	5.74 ± 2.85
Five	Control (n = 16)	Pre-test	4.13 ± 2.05
		Post-test	3.77 ± 2.11
	Exp. (n = 15)	Pre-test	2.67 ± 1.18
		Post-test	4.07 ± 2.32

4 Discussion

The "A me gli occhi" project stems from the hypothesis that training visuospatial abilities in primary school students would have a positive impact on their school performance in math. After taking a pre-test, students in the experimental group have undergone twenty game based training sessions focused on visuospatial abilities. At the end of the project, a post-test was taken and results have been compared to the pre-test.

In the written operations section, an initial disadvantage of the grade 5 experimental class compared to the control one was measured at the pre-test. Such a disadvantage has disappeared at the post-test, when the experimental class got better results than the control class. In the ordering task, as well as in the word problems, both the experimental classes got better scores than the control classes at the post-test and the improvement in scores between the pre-test and the post-test were significant only in the experimental group. This shows that the visuospatial abilities training has actually shown some effect on the experimental groups.

Moreover, all the experimental groups show an increasing trend in the scores in all the sections of the post-test when compared to the pre-test. The same cannot be said for the control groups, which shows some improvement in some areas, but not in all of them. The task of identifying the biggest between two numbers did not get any significant effect, but the trends were very interesting since both experimental classes show an increase of the average score and a decrease of the standard deviation.

The "AC-MT 6-11" math test [22] that has been used was available for the different grades and could be given and marked by class teachers, but it showed some drawbacks. For example, word problems have caused some difficulty in their interpretation from a linguistic point of view; some drawings in the ordering part of the test suggested a different order from what was asked, causing some confusion. These issues could be addressed in future experiments.

As far as the participants' enjoyment in the activities, there has been a very strong interest in the Minecraft contest. A possible interpretation is that the students become more involved in those activities that require them to be more creative As a matter of facts, Minecraft, being a sandbox game [25], involves its users not only as players of a readymade game, but also, in a more creative manner, as builders of their own digital experience. This will have to be taken into consideration in future projects.

In all activities, the students' participation and interest has been very high. In general, all the suggested activities have been accepted with enthusiasm as confirmed by the data collected from the questionnaires on game enjoyment; nevertheless, some games have been liked more than others, as well as some version of the same game has been played with more participation than different versions. The experimenters have been rather flexible, offering, at each session, more than one game, always paying attention to maintain the same kind of tasks as far as the visuospatial abilities were concerned.

5 Conclusions

In this paper the "A me gli occhi" project has been presented. The main goal of this project was to verify if a specific training of visuospatial abilities would have an impact on math school results in students from grades four and five of an Italian primary school. During the project, twenty game based training sessions have been organized, of approximately one hour each, from November 2016 to May 2017. Games were mostly digital apps on an Android tablet, but a computer based game and some concrete gaming activities have also been carried out. Each student played individually with one tablet, but peer support was encouraged. A standardized math test has been given as a pre-test in November and as a post-test in May to the two experimental classes and the two control classes. Results from the pre and post-test have then been analysed.

Data analysis has shown that there has been an increase in the performance between the beginning of the project and its end. Such improvement is more relevant in the experimental group than in the control one and it is statistically significant in some areas of the standardized math test that has been used. This is in line with our hypothesis, according to which having good visuospatial abilities has a positive influence on school achievements in mathematics.

The game based training activities, which have been carried out in one school year, have involved the researchers working directly with the students. In future works, in order to allow a much wider participation, class teachers will have to be actively involved. This will allow organizing a specific training of the students' visuospatial activities in a more flexible manner, enriching the lessons every time it would be relevant. Future studies, lasting for a longer period, may also consider taking into account a closer monitoring of the students' performance and attitudes.

Participants have showed a strong involvement in all the game activities and have participated with interest in the project. Nevertheless, the involvement and interest measured have been higher in those games that required a more active participation from the children, as it has been the case for the Minecraft contest, Origami, etc. Future works may focus more on students' creativity, promoting not only game playing but also their creation, following a constructionist approach [26], as was the case for the houses built inside Minecraft during the contest.

Acknowledgments. The authors would like to thank Carlo Chiorri and Ottavia Epifania for their help in the statistical analysis of the collected data. Micòl Arena and Beatrice Crippa for their valuable help working in class with the researcher and the students during the training sessions. Last but not least, all the involved class teachers and students for their heartfelt participation.

Appendix: List of Digital Games Used in the Project

The following is a list of the games that have been used for the training sessions in the project. “In Your Eyes” has been implemented for Windows or iOs computers. All the others are Android Apps.

In Your Eyes. The game asks the player to recognize the picture reproducing a scene as it would be seen from a different point of view. The game has been developed by the Italian CNR-ITD. Available also for use also with Oculus Rift in its SDK2 version.

Memory training. The app offers four different games, all aimed at improving the visual memory of the player.

Jigsaw puzzles. Several free apps have been used, with the possibility to choose different pictures or drawings, define both the number of pieces, and rotation. During the training, students started with the simplest puzzles, incrementing the number of pieces and inserting rotation at they got more experienced.

Studio 4. A collection of games supporting recognition and association of colours, forms, and patterns; comparison of quantities and measurements; spatial orientation; eye-hand coordination; counting and recognition of quantities up to 10. The game has not been used only in grade 4 since it has been judged too simple for grade 5.

Tangram. The classical Tangram game offers a wide number of compositions to be reproduces with the classical tangram pieces.

Pipe Twister. The player, by tapping pipes, has to make a pipeline to connect coloured water flows from a source to the matching output pipe.

Origami. It is possible to find several instructions free on Android that have to be followed by folding real paper. Furthermore, a few digital origami game apps on Android exist, but they have resulted much less engaging than the real world activity.

Tetris. The classical Tetris game.

Mekorama. The game features a nice little robot that has to go from a starting point to a destination by finding its path through a complex 3D structure.

Minecraft. The famous sand box game, used only as a peaceful construction tool starting from a flat land.

Mazes. Several free Maze games on Android are available, with different levels of difficulty.

Evo Explorer. Similar to Mekorama, it involves moving a character from one spot to another, but with the characteristics that what cannot be seen does not exist. It plays with perspective creating new routes as the structure is turned around and the point of view on it changes.

Piko's Blocks. A series of exercises where a structure made of blocks is given and it has to be copied into a new structure. Sometimes assail symmetry is also used: the new structure has to be the reflection of the old one.

References

1. Wai, J., Lubinski, D., Benbow, C.P.: Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. J. Educ. Psychol. **101** (4), 817–835 (2009)
2. Sinclair, N., Bruce, C.D.: New opportunities in geometry education at the primary school. ZDM **47**(3), 319–329 (2015)
3. Carlisle, D., Tyson, J., Nieswandt, M.: Fostering spatial skill acquisition by general chemistry students. Chem. Educ. Res. Pract. **16**(3), 478–517 (2015)
4. Uttal, D.H., Cohen, C.A.: Spatial thinking and STEM education: when, why and how. Psychol. Learn. Motiv. **57**(2), 147–181 (2012)
5. Uttal, D.H., Meadow, N.G., Tipton, E., Hand, L.L., Alden, A.R., Warren, C., Newcombe, N. S.: The malleability of spatial skills: a meta-analysis of training studies. Psychol. Bull. **139** (2), 352 (2013)
6. Newcombe, N.S., Frick, A.: Early education for spatial intelligence: Why, what, and how. Mind Brain Educ. **4**(3), 102–111 (2010)
7. Bottino, R.M., Ott, M., Tavella, M.: Serious gaming at school: reflections on students' performance. In: Gamification: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications, vol. 314. IGI Global (2015)
8. Bottino, R.M., Ott, M.: Mind games, reasoning skills, and the primary school curriculum: hints from a field experiment. Learn. Media Technol. **31**(4), 359–375 (2006). Taylor and Francis
9. Bednarz, R.S., Lee, J.: The components of spatial thinking: empirical evidence. Proc. Soc. Behav. Sci. **21**, 103–107 (2011)
10. Baddeley, A.D., Hitch, G.: Working memory. Psychol. Learn. Motiv. **8**, 47–89 (1974)
11. Mulligan, J.: Looking within and beyond the geometry curriculum: connecting spatial reasoning to mathematics learning. ZDM **47**(3), 511–517 (2015)
12. Lowrie, T.: Digital games, mathematics and visuospatial reasoning. In: Lowrie, T., Jorgensen (Zevenbergen), R. (eds.) Digital Games and Mathematics Learning. MEDE, vol. 4, pp. 71–92. Springer, Dordrecht (2015). https://doi.org/10.1007/978-94-017-9517-3_5
13. Bottino, R.M., Ott, M., Tavella, M.: Serious gaming at school: reflections on students' performance, engagement and motivation. Int. J. Game-Based Learn. (IJGBL) **4**(1), 21–36 (2014). https://doi.org/10.4018/IJGBL.2014010102. IGI Global
14. Bottino, R.M., Ott, M., Tavella, M.: Children's performance with digital mind games and evidence for learning behaviour. In: Lytras, M.D., Ruan, D., Tennyson, R.D., Ordonez De Pablos, P., García Peñalvo, F.J., Rusu, L. (eds.) WSKS 2011. CCIS, vol. 278, pp. 235–243. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-35879-1_28. ISBN 978-3-642-35878-4

15. Bottino, R.M., Ferlino, L., Ott, M., Tavella, M.: Developing strategic and reasoning abilities with computer games at primary school level. Comput. Educ. **49**(4), 1272–1286 (2007). Elsevier
16. Mathewson, J.H.: Visual-spatial thinking: An aspect of science overlooked by educators. Sci. Educ. **83**, 33–54 (1999)
17. Surtees, A.D.R., Apperly, I.A., Samson, D.: The use of embodied self-rotation for visual and spatial perspective-taking. Front. Hum. Neurosci. **7**, 698 (2013). https://doi.org/10.3389/fnhum.2013.00698
18. Freina, L., Bottino, R., Tavella, M., Dagnino, F.: Immersion's impact on performance in a spatial reasoning task. In: Bottino, R., Jeuring, J., Veltkamp, R.C. (eds.) GALA 2016. LNCS, vol. 10056, pp. 211–220. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-50182-6_19
19. Meyer, M.L., Salimpoor, V.N., Wu, S.S., Geary, D.C., Menon, V.: Differential contribution of specific working memory components to mathematics achievement in 2nd and 3rd graders. Learn. Individ. Differ. **20**(2), 101–109 (2010)
20. Piaget, J., Inhelder, B.: The Child's Conception of Space. Routledge and Kegan Paul, London (1956). Langdon, F.J., Lunzer, J.L. (Trans.)
21. Newcombe, N.: The development of spatial perspective taking. Adv. Child Dev. Behav. **22**, 203–247 (1989)
22. Cornoldi, C., Cornoldi, C., Lucangeli, D., Bellina, M.: AC-MT 6-11. Test di valutazione delle abilità di calcolo e soluzione dei problemi. Gruppo MT. Con CD-ROM. Edizioni Erickson (2012)
23. Freina, L., Canessa, A.: Immersive vs desktop virtual reality in game based learning. In: Proceedings of the 9th European Conference on Games Based Learning (ECGBL), Steinkjer, Norway, October 8–9, 2015, pp. 195–202 (2015)
24. Alloway, T.P., Passolunghi, M.C.: The relationship between working memory, IQ, and mathematical skills in children. Learn. Individ. Differ. **21**(1), 133–137 (2011)
25. Ekaputra, G., Lim, C., Eng, K.I.: Minecraft: A game as an education and scientific learning tool. In: 2013 Information Systems International Conference (ISICO) (2013)
26. Kafai, Y.B.: Playing and making games for learning: Instructionist and constructionist perspectives for game studies. Games Cult. **1**(1), 36–40 (2006)

Games for Health and Special Children

BeeSmart: A Gesture-Based Videogame to Support Literacy and Eye-Hand Coordination of Children with Down Syndrome

Veronica Lizeth Amado Sanchez[1], Oscar Iván Islas Cruz[1], Edgar Armando Ahumada Solorza[1], Iván Alejandro Encinas Monroy[1], Karina Caro[2(✉)], and Luis A. Castro[1]

[1] Sonora Institute of Technology (ITSON), Ciudad Obregon, Sonora, Mexico
veronicaamado4@gmail.com, oislas03@gmail.com, eahumadasolorza@gmail.com, ivanencinasm@gmail.com, luis.castro@acm.org

[2] Drexel University, Philadelphia, PA, USA
karinacaro@drexel.edu

Abstract. Children with Down Syndrome (DS) have deficits in eye-hand coordination skills. These deficits impact activities of daily living (ADL) and children's independence. In addition, there is a substantial relationship between eye-hand coordination and the acquisition of literacy skills. In this paper, we present the design process of a gesture-based videogame to support eye-hand coordination and literacy skills. First, we conducted a contextual study in an educational institute for children with DS. Next, using the contextual study results, several design solutions were generated and iterated using low-fidelity prototypes. Finally, we designed and developed *BeeSmart*, a videogame based on the Troncoso's method for literacy, wherein children use their fingers to draw around pictograms and words on a screen. Finally, we present a usability evaluation of *BeeSmart* to show the potential for supporting children with DS.

Keywords: Videogame · Literacy · Eye-hand coordination · Down syndrome

1 Introduction

Down Syndrome (DS) is a disorder caused by an extra 21st chromosome and it occurs in 1 about 700 births [1]. Children with DS experience delayed motor development and difficulty performing functional motor ability. Specifically, children with DS have deficits in eye-hand coordination, laterality, visual-motor control, and visual spatial perception [2]. Deficits in eye-hand coordination[1] impact negatively on activities of daily living, decreasing the possibility of achieving children's independence.

[1] Eye-hand coordination is the ability to integrate both visual and motor abilities to accomplish a goal (e.g., reaching a ball). It involves using hands/fingers in order to reach a visual target [8].

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 43–53, 2017.
https://doi.org/10.1007/978-3-319-71940-5_4

Literacy, the ability to read and write, is an activity based on written and spelling systems [3]. Eye-hand coordination is one of the five components that contribute to the quality of handwriting: kinesthesia, motor planning, eye-hand coordination, visual-motor integration, and in-hand manipulation. Deficits in eye-hand coordination could impact the acquisition of literacy skills [4]. Due to the specific characteristics of children with DS, they might need a specific approach to learn to read and write [3]. In [5], a literacy method is proposed to teach children with DS: the Troncoso's method. The method is designed to recognize words as a whole, instead of breaking words into letters. The philosophy of the Troncoso's method is to recognize the meaning of written symbols (i.e., words) by association. Children recognize words making relationships between the written word and its associated pictogram [5].

Several works have demonstrated that an intervention formed by different activities to support eye-hand coordination (e.g., drawing following visual patterns) have a positive impact on the eye-hand coordination of children with DS [4, 6]. Serious Games, digital games with an agenda of educational or therapeutic design, beyond entertainment, are a promising teaching and therapeutic tool to support this population due to its interactivity and engaging and immersive activities [7]. Particularly, serious gesture-based videogames, videogames that are controlled using gestures without wearing additional aides (e.g., body markers) might have the potential to support eye-hand coordination, as they can track hand movements to execute some actions on the screen. In addition, considering that play is an important facilitator and an important means for the development of cognitive and motor coordination skills [8], we believe that this technology has potential to support the literacy process. It can be based on literacy methods such as [5], combined with gamification techniques such as points and rewards [9] for providing adequate solutions for this population.

1.1 Technology to Support Eye-Hand Coordination in DS

There are a few studies related to eye-hand coordination and the use of technology that enables users to use their hands to follow visual patterns [10, 11]. For example, there are studies aimed at investigating the use of tablet-based technology to support handwriting skills [10] of neurotypical 3-year-old children and works of exergames using Kinect-based games for supporting eye-hand coordination of children with motor coordination problems such as autism [11]. These works show that Tablet and Kinect based technology could support eye-hand coordination skills of neurotypical children and children with motor problems such as cerebral palsy and autism.

Other works have investigated the use of free-hand interaction using gesture tracking sensors such as Leap Motion (http://leapmotion.com) to support hand rehabilitation of patients with stroke or trauma [12, 13]. In [12], authors present a framework for the gamification of hand therapies, particularly on the wrist joint range of motion (ROM) therapies of individuals with trauma or stroke. The framework consists of an adaptive therapy-driven 3D environment augmented with a motion-based natural user interface. To evaluate the framework, several games were implemented using the Leap motion sensor to track user's hand movements. The games were evaluated with 5 neurotypical children, resulting in highly-motivated children when playing games.

1.2 Technology to Support Literacy for DS

Several studies have explored the use of technology to support literacy of children with developmental disabilities, such as significant developmental delays [14]. Particularly, there a few studies investigating the use of technology to support literacy of children with DS [15, 16]. For example, the work of [15] presents a tangible user interface (TUI) system based on RFID to support literacy for children with DS. The system focuses on the integration of RFID tags in 3D printed objects and low-cost toys. Children relate the written word with its associated 3D printed object. The study of [16] presents HATLE, a system based on mobile computing using tablet devices, multimedia design, and computer speech recognition for improving reading and writing abilities in Spanish, through speech and drawing activities.

Although promising, some of the aforementioned works aim at supporting populations with motor problems derived from cerebral palsy, autism, stroke or trauma. Thus, open questions remain as to how these results can extend to other population such as children with DS. As discussed, not many works are aimed at supporting children with DS. In this paper, we present the design and evaluation of a gesture-based videogame to support eye-hand coordination and literacy skills of children with DS. First, we present a contextual study with the goal of understanding the challenges and problems faced by children with DS in terms of motor coordination and literacy. Next, we show how we used the contextual study findings to design and develop a gesture-based videogame to support eye-hand coordination and literacy. Finally, we present a usability evaluation of the gesture-based videogame, showing the potential of this technology supporting eye-hand coordination and literacy skills.

2 Designing Gesture-Based Videogames for Children with DS

To guide our research process, we used the design thinking methodology, fundamentally consisting of five main phases: (1) Empathize, (2) Define, (3) Ideate, (4) Prototype and (5) Test [17]. For phases 1 and 2, we conducted a contextual study to gain an understanding of children with DS and their intervention program. For the phases 3 and 4 and using the contextual study results, we conducted several design sessions to generate initial potential design solutions. Using low-fidelity prototypes, we iterated the design solutions. Reflecting on the contextual study findings, we selected and refined the best design solution. For phase 5, we tested the design solution using a medium-fidelity prototype through a focus group with three psychotherapists who work with children with DS. Finally, we used the focus group results to improve our design solution and we implemented it as a high-fidelity prototype.

2.1 Contextual Study

A contextual study was conducted for 4 weeks in an institute for intervention and education for children with DS, partially funded by a non-profit NGO, located in Northwest, Mexico. The institute serves the needs of 15-20 children with DS, providing cognitive, motor, and sensorial therapies. It is run by 15 volunteers (neuropsychologists, physiotherapists, etc.) and family members who help with certain tasks.

Data collection. We conducted 9 semi-structured interviews; 4 with psychotherapists (AVG age = 22.5 (±1), one male), 1 with a neuropsychologist (age = 40, male) and 4 with parents of children with DS (AVG age = 45.7 (±3.3), all females). The interview guide included topics as the types of interventions they provide to children with DS, the common problems they face when providing therapy, the characteristics of this population, among others. The average interviews lasted 34:37 min (±13:48 min). All the interviews were audio recorded and transcribed for posterior analysis. Consent forms were collected from all informants. Also, 11 therapies were observed, including cognitive (n = 5), motor (n = 5), and emotional expression (n = 1) therapies. The average therapy session lasted 18:44 min (±09:56 min). 4 researchers observed children during therapies using the framework for analyzing observational data [18], often used in interaction design [19].

Data analysis. We used qualitative techniques to derive grounded theory such as open and axial coding [20] and affinity diagramming [21]. Quotes obtained from interviews and observation-notes were grouped for uncovering potential emerging themes related to characteristics of children with DS, intervention program, problems therapists faced during therapies, and the strategies they use.

Results. Our results focus mainly on the manner in which therapies are carried out, −normally referred to as intervention programs−emphasizing how technology could support these activities. The intervention program is tailored to each child's needs. During the motor therapies, children perform activities to support eye-hand coordination, laterality, and visual spatiality such as following visual patterns or trajectories and locating objects in a specific space. As one of the main priorities in the development of children with DS is independence, the acquisition of reading and writing skills plays a very important role. The institute uses the Troncoso's method [5] to teach children to read and write, in which therapists use cards with the written symbol and its associated pictogram to teach children to recognize the words as a whole, associating the written words with their representative pictograms. Once children are able to identify some vocabulary (written words and pictograms), therapists proceed to break down the words into letters or group of letters (i.e., syllables). At the beginning, to encourage children to recognize different words and because of the strong relationship between eye-hand coordination and writing skills, therapists ask children to use their fingers to trace the written word and its associated pictogram. In this way, children work with literacy and eye-hand coordination skills. Therapists constantly use positive verbal reinforcement for maintaining children engaged with the therapy's activities (e.g., *Good Job! You are working so well, today!*).

Design considerations. Reflecting on the contextual study findings, the following design considerations were obtained for designing technology to support literacy and eye-hand coordination of children with DS: (a) the solution must imitate the Troncoso's method [5], allowing children to work with written words and their pictograms; (b) it should support eye-hand coordination activities, imitating the way therapists work with children during the therapies; and, (c) the activities should be gamified using visual and positive verbal reinforcements to promote children's motivation.

2.2 Design Sessions and Validation

Iterative design. Using the design considerations, three iterative design sessions were conducted using the sketching technique [22], where the research team generated low fidelity prototypes of initial design solutions. The general idea was using the forefinger to draw a written word following a dotted-path or using fingers to drag and drop a group of letters to order them to form a word. The solution consists of *BeeSmart*, a serious game with four activities: (A1) drawing a pictogram (e.g., a tree), following a dotted-path, (A2), drawing the word associated with the pictogram in A1, (A3) ordering syllables to form a word, (A4) ordering letters to form a word. Interaction is thru air gestures with fingers, mainly the forefinger. Children achieve points for each activity.

Validation through a focus group. To validate the design of our solution, we conducted a focus group (01:05 h) with three psychotherapists (AVG age = 36.33 (±10.97), 1 male). During the focus group, a video with the main characteristics of the solution was presented to the participants. After the video presentation, several design aspects were discussed. For example, the mechanics of the activities, the interaction mode, the visual aspect of the game (e.g., avatars and colors), among others. The focus group yielded results in terms of (a) activities, (b) reinforcement and rewards, and (c) interaction mode. For activities, the most important activities to start teaching children to read and write relate to (1) drawing a pictogram, and (2) drawing the associated word. Suggestions included having different topics to teach several words, showing both visual and audible cues, using upper/lower letter cases, and the like. On the other hand, for reinforcements and rewards, suggestions included giving a piece of a puzzle to children instead of achieving points, positive visual or verbal reinforcements at the end of the activity (e.g., a happy face). Finally, for the interaction mode, psychotherapists agreed in using air gestures, particularly the forefinger, since this could help children with both eye-hand coordination and visual-spatial perception.

3 BeeSmart: A Gesture-Based Videogame for Children with DS

Taking the results of the focus group as a departure point, we designed *BeeSmart,* a game to support literacy and eye-hand coordination of children with DS, based in the Troncoso's method for literacy.

Game mechanics. *BeeSmart* demands children to draw different words and their pictograms using their forefinger. At the beginning, the therapist selects the topic (e.g., shapes, animals) (Fig. 1a). Once the topic is selected, the therapist selects the specific word to be shown (e.g., square, circle), and one of the two activities: draw the pictogram or draw the word. In both activities, the pictogram, the written word, and its pronunciation are shown and played.

Story game. The story behind *BeeSmart* is a bee that needs help to learn new words. The story is shown to children as a short animated video at the beginning of the game, inviting children to help it learn new words.

Verbal and visual prompts and reinforcements. When an activity starts, the game plays the pronunciation of the word. The pattern to follow for drawing the pictogram or the word is shown using dashed-lines (Fig. 1b). Additionally, each time children perform an activity successfully, positive verbal reinforcements (e.g., *excellent job!*) are played and visual positive reinforcements are shown (e.g., a bee with a happy face). *BeeSmart* also gives a piece of a puzzle each time children complete an activity. The puzzle is related to the word that the player is working (e.g., if the word is “square”, the puzzle shows a square window). Finally, if more than a minute passes and the player is not drawing the shape, the game provides verbal prompts to motivate children to continue playing (e.g., *you can do it! Try again!*).

Profile and progress. To maintain a record of the players’ progress and performance, each player has a profile. A profile is created for each new player. The profile records each activity´s duration and number of trials to complete an activity. In addition, children’s drawings are saved to show therapists their progress (Fig. 1c).

Fig. 1. Screenshots of the final design of *BeeSmart*. (Color figure online)

Game configuration. Since the goal of *BeeSmart* is supporting therapists in literacy and eye-hand coordination skills, the game configuration is managed by them (e.g., select the player, the topic, and the activity).

Gesture tracking. Players’ hand gestures are tracked using the Intel® RealSense™ SR300 Sensor. The tracking is activated once the therapist starts a new activity. Intel® RealSense™ Sensor tracks children’ forefinger, and it enables children to perform the activity either with the right or the left hand. In *BeeSmart*, the Intel® RealSense™ operates mainly with a latency between 11 and 14 milliseconds (ms), with certain peaks at 6 ms (min.) and 18 ms (max.). The Intel® RealSense™ latency was calculated using the Stopwatch class of C# language. Stopwatch started counting when the therapist started an activity and stopped when a new frame within the activity was updated.

Implementation details. *BeeSmart* was implemented using C# language and Unity game engine development platform (http://unity3d.com). *BeeSmart* follows a one-tier architecture with three layers (i.e., presentation, business, and data layer) in a single software package. All data is stored on the local system (i.e., using a PC with an Intel Core i5 5th generation processor and 8 GB RAM). In addition, *BeeSmart* follows the model-view-controlled (MVC) architectural pattern, dividing the game into three interconnected parts and separating internal representation of information from the ways information is presented to the player. The model manages the data obtained from the controller (i.e., stores data that is retrieved according to commands from the controller and displayed in the view), the logic and rules of the game. The view generates the game visual interfaces to the player based on changes in the model. The controller accepts input data from the Intel® RealSense™ sensor and sends commands to the model to update the model's state. To establish communication between the Intel® RealSense™ Sensor and the controller of the game, we used Intel® RealSense™ SDK R2. In addition, to validate the players' drawings, key points were located in the pattern to follow in the pictogram or word. These key points are colored in yellow (Fig. 1b), once a player draws over a key point, the point changes to red (Fig. 1c). Also, a sound is played as feedback once the player draws over a key point.

4 Usability Evaluation

4.1 Methods

Participants. 12 children with DS between 6 and 13 years old (AVG age = 10.2 years, ±2.9 years, 6 males) played *BeeSmart* during one game session (Fig. 2). None of the participants had previous experience using fingers tracking sensors.

Fig. 2. A participant and her therapist during the usability evaluation

Data collection. Children were allocated in groups of three i.e., three children per session. Each child was asked to draw pictograms for two consecutive rounds (R1 and R2) with 5 shapes suggested by therapists: square, circle, rectangle, triangle, and star. All game sessions were video recorded. Additionally, *BeeSmart* recorded each activity's duration and the number of attempts per participant. To measure the game experience of children with DS playing *BeeSmart*, we used the In-game module of the Game Experience Questionnaire (GEQ) [23]. The GEQ was directly answered by 3 participants, and the rest were answered by the proxy (therapist). At the end of each session, a short interview was conducted with the therapist (AVG duration = 06:17 min, ±00:10 min). Topics included perceptions on children interactions with the game, main problems observed, and likely improvements.

Data analysis. We used descriptive statistics to analyze player game logs and the GEQ questionnaire. For the semi-structured interviews, we used qualitative techniques such as open and axial coding [20] and affinity diagraming [21] to uncover emergent themes related to the game experience of children using *BeeSmart*.

4.2 Results

Participants' performance and game experience perception. Table 1 shows the results of the performance of our participants. Each session lasted about 25 min. From the 12 participants, 2 were unable to complete the activity for each geometric shape in R2. Thus, these 2 participants had zero attempts and zero seconds in each shape's duration as they did not perform R2. Following Table 1, on average, each participant made 5.07 attempts to complete each of the five tasks during R1; and 1.67 attempts during R2. The geometric shape that, on average, needed more attempts in R1 was the circle, followed by the star. In comparison, the triangle seemed to be much less complex. For R2, all participants performed better on average i.e., fewer attempts needed to complete the task. On the other hand, taking into account the time to complete each of the tasks (i.e., drawing each shape), the average attempt's duration was slightly more than 30 s. In particular, the square was the shape that took longer to complete, while the circle and the star were drawn faster than average.

Table 1. Descriptive results of the participants' attempts per round

	Attempts per shape (#)		Time per attempt per shape (sec.)	
	R1 (n = 12)	R2 (n = 10)	R1 (n = 12)	R2 (n = 10)
Square	5.25 (6.21)	**1.00 (0.60)**	41.69 (9.49)	43.42 (28.50)
Circle	9.67 (13.06)	1.17 (0.94)	**24.38 (7.43)**	27.24 (17.92)
Triangle	**1.92 (2.27)**	1.33 (1.82)	34.21 (17.60)	24.92 (13.94)
Rectangle	2.42 (1.92)	1.42 (1.38)	36.15 (27.09)	26.49 (16.50)
Star	6.08 (8.86)	3.42 (5.79)	28.04 (18.22)	**24.32 (15.90)**
AVG	5.07 (3.13)	1.67 (0.99)	32.89 (6.81)	29.28 (7.99)

As for the GEQ (see Table 2), the overall game experience was perceived as positive. In particular, the scores for competence, immersion and positive affect dimensions are above 3 points, indicating that children felt skillful, they were interested, happy or satisfied, and they felt good when they were playing the game. In comparison, the tension and negative affect had scores below 1, indicating that children did not feel frustrated, bored or irritable when playing the game. Finally, the challenge dimension had a score of 2.71, indicating that at least half of the participants felt challenged.

Table 2. Average GEQ results and standard deviations by dimensions (Min = 0, Max = 4)

Competence	Immersion	Flow	Tension	Challenge	(–) Affect	(+) Affect
3.29 (1.54)	**3.79 (0.33)**	0.83 (1.03)	0.21 (0.4)	2.71 (1.45)	0.13 (0.23)	**3.96 (0.14)**

Therapist perception. During the post-session interviews, the therapist shared with us several observations and assessment of a qualitative nature. This proved extremely valuable since she knows the children and works with them on a weekly basis. Regarding the performance of children, the therapist expressed that the circle and the star were the most difficult shapes at the beginning because these shapes require more eye-hand coordination, since they have more key points to follow: "*I think that [BeeSmart] can support eye-hand coordination. In this case for literacy skills, children need to measure spaces, for instance, in their notebooks, where they are going to trace, they need to pay attention to where the trace is, focus, and [even] memorize some traces. We saw that in the second round they made it much easier. The same happens when we are learning to trace letters. We struggle with the first one but then the next ones are simpler.*" In addition, she provided the example of one of the children with whom they had begun working with letter tracing: *"… with P6 hardly… we have begun with the tracing, and we can hardly keep him focused for longer. For me, it was surprising that he could last 5 min., even though he didn't make it [accomplish the task]… he wanted to keep trying at least in the first 3 shapes. I think it [BeeSmart] was something that caught his attention and encourage him to keep on trying.*"

On the other hand, regarding game experience, the therapist agreed that all children seemed very excited to play *BeeSmart*: "*Today, I saw that all four children were very happy when finishing, and seeing the little bee. It was very pleasing for them… when they watched the prize figure, I saw them very happy because they were images that they related to. They quickly identified a ball, a pizza, a door, without even being finished, and that made them feel, like skillful, satisfied… I saw them very happy.*"

Finally, regarding the interaction mode, all participants identified that the movements of their forefinger have the control in the activities available in *BeeSmart*, which was mentioned by the therapist: "*The children quickly understood that their hand was going to move the pencil that is in front of them, then they are focused in what they are seeing with the movements of their hands.*" More, specifically, since *BeeSmart* enables children to use either their right or left hand, children used the hand that they felt more comfortable with. This feature was highlighted by the therapist: "*It helps with eye-hand coordination. I saw that with P10, since she's left-handed, and she did it easily. She didn't actually struggle to discriminate which hand she was going to use… and that's one step forward in literacy, since children must decide, according to their skills, with which hand they will write.*"

BeeSmart supported eye-hand coordination and literacy in our participants, promoting activity tolerance, visual attention, and visual spatial perception. Also, positive visual and verbal reinforcement were appropriate for our participants, since all were motivated and excited every time they received the positive reinforcements.

5 Conclusions and Future Work

We presented the design and evaluation of *BeeSmart*, a videogame to support literacy and eye-hand coordination skills of children with DS. The usability evaluation showed that following our design considerations generated from the contextual study resulted in a positive gaming experience for our participants. Moreover, the usability evaluation showed several improvements to the design of *BeeSmart*. For example, more complex pictograms were suggested by the therapist for children with a higher cognitive level (e.g., a house). The game should enable therapists to tailor activities per child in terms of pictograms, visual and verbal prompts, or number and kind of positive verbal reinforcements. As future work, we will evaluate the efficacy of *BeeSmart* through a long-term evaluation study.

Acknowledgements. We thank CEART DOWN AC for the continuous support.

References

1. Davis, A.S.: Children with down syndrome: Implications for assessment and intervention in the school. Sch. Psychol. Q. **23**, 271–281 (2008)
2. Spanò, M., Mercuri, E., Randò, T., Pantò, T., Gagliano, A., Henderson, S., Guzzetta, F.: Motor and perceptual-motor competence in children with Down syndrome: variation in performance with age. Eur. J. Paediatr. Neurol. **3**, 7–14 (1999)
3. Badillo Jiménez, V.T.: Using picture communications symbols (PCS) for teaching to read and write children with down´s syndrome. Int. J. Educ. Sci. Res. **4**, 1–8 (2014)
4. Varuzza, C., De Rose, P., Vicari, S., Menghini, D.: Writing abilities in intellectual disabilities: a comparison between down and williams syndrome. Res. Dev. Disabil. **37**, 135–142 (2015)
5. Troncoso, M.V., Del Cerro, M.M.: Síndrome de Down: Lectura Y Escritura. Fundación Iberoamericana Down21 (2009)
6. Lersilp, S., Putthinoi, S., Panyo, K.: Fine motor activities program to promote fine motor skills in a case study of down's syndrome. Glob. J. Health Sci. **8**, 60 (2016)
7. Thompson, D., Baranowski, T., Buday, R., Baranowski, J., Thompson, V., Jago, R., Griffith, M.J.: Serious video games for health: how behavioral science guided the development of a serious video game. Simul. Gaming **41**, 587–606 (2010)
8. Gallahue, D.L., Ozmun, J.C.: Understanding Motor Development Infants, Children, Adolescents, Adults. McGraw-Hill Humanities, Social Sciences & World Languages (1998)
9. Deterding, S., Dixon, D.: From game design elements to gamefulness: defining "Gamification". In: MindTrek 2011, pp. 9–15 (2011)
10. de Diego-Cottinelli, A., Barros, B.: TRAZO: a tool to acquire handwriting skills using tablet-PC devices. In: International Conference on Interaction Design and Children, pp. 278–281. ACM, Barcelona (2010)

11. Caro, K., Tentori, M., Martinez-Garcia, A.I., Alvelais, M.: Using the FroggyBobby exergame to support eye-body coordination development of children with severe autism. Int. J. Hum Comput. Stud. **105**, 12–27 (2017)
12. Afyouni, I., Qamar, A.M., Hussain, S.O., Ur Rehman, F., Sadiq, B., Murad, A.: Motion-based serious games for hand assistive rehabilitation. In: International Conference on Intelligent User Interfaces Companion, pp. 133–136 (2017)
13. Iosa, M., Morone, G., Fusco, A., Castagnoli, M., Fusco, F.R., Pratesi, L., Paolucci, S.: Leap motion controlled videogame-based therapy for rehabilitation of elderly patients with subacute stroke: a feasibility pilot study. Top. Stroke Rehabil. **22**, 306–316 (2015)
14. Chai, Z., Vail, C.O., Ayres, K.M.: Using an iPad application to promote early literacy development in young children with disabilities. J. Spec. Educ. **48**, 268–278 (2015)
15. Jadán-Guerrero, J., Guerrero, L., López, G., Cáliz, D., Bravo, J.: Creating TUIS using RFID sensors—a case study based on the literacy process of children with down syndrome. Sensors **15**, 14845–14863 (2015)
16. Felix, V.G., Mena, L.J., Ostos, R., Maestre, G.E.: A pilot study of the use of emerging computer technologies to improve the effectiveness of reading and writing therapies in children with down syndrome. Br. J. Educ. Technol. **48**, 611–624 (2017)
17. Johansson-Sköldberg, U., Woodilla, J., Çetinkaya, M.: Design thinking: past, present and possible futures. Creat. Innov. Manag. **22**, 121–146 (2013)
18. Robson, C.: Real World Research: A Resource for Social Scientists and Practitioner-Researchers. Blackwell, Maldren (2002)
19. Rogers, Y., Sharp, H., Preece, J.: Interaction Design: Beyond Human Computer Interaction. Wiley (2011)
20. Strauss, A., Corbin, J.: Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. SAGE Publications (1998)
21. Holtzblatt, K., Wendell, J.B., Wood, S.: Rapid Contextual Design: A How-to Guide to Key Techniques for User-Centered Design (2005)
22. Greenberg, S., Carpendale, S., Marquardt, N., Buxton, B.: Sketching User Experiences. Morgan Kaufmann, Elsevier (2012)
23. Ijsselsteijn, W., Poels, K., de Kort, Y.A.: The game experience questionnaire: development of a self-report measure to assess player experiences of digital games. TU Eindhoven, The Netherlands (2008)

Sinbad and the Magic Cure: A Serious Game for Children with ASD and Auditory Hypersensitivity

Hanan Makki Zakari[1(✉)], Matthieu Poyade[1], and David Simmons[2]

[1] The School of Simulation and Visualisation,
The Glasgow School of Art, Glasgow, UK
H.Zakari1@student.gsa.ac.uk, M.Poyade@gsa.ac.uk
[2] The School of Psychology, University of Glasgow, Glasgow, UK
David.Simmons@glasgow.ac.uk

Abstract. Serious games (SG) are part of a range of technologies that can be used for children with Autism Spectrum Disorder (ASD). Typically, these games target communication skills, social interaction and speech. There are, however, few SGs for autistic children which have the aim of helping to moderate sensory hypersensitivity. This paper presents an SG which aims to target this area. A set of critical sounds was heuristically determined, and implemented in an interactive SG for reducing auditory hypersensitivity (i.e. over-sensitivity to environmental sounds) in children with ASD. The game, *Sinbad and the Magic Cure*, is designed for Android devices, and is intended for children aged 8–11.

Keywords: Serious games · Game-based therapy · Autism Spectrum Disorder Moderating sensory hypersensitivity · Auditory hypersensitivity · Hyperacusis

1 Introduction

Autism Spectrum Disorder (ASD) has two main groups of symptoms: issues related to social communication and social interaction; and repetitive patterns of behavior, interests or activities [1]. The number of children diagnosed with ASD has continued to grow in recent decades, with estimates as high as one in 68 in the US [2] and one in 100 in the UK [3] suffering from this condition. Up to 90% of autistic children have auditory, visual and/or tactile hypersensitivity [4], and SGs may provide significant assistance in moderating these hypersensitivities. In this paper, we present an SG for autistic children with auditory hypersensitivity. We have designed an adventure SG, *Sinbad and the Magic Cure*, to be played on mobile android devices, and investigate the following research questions:

- What sounds appeared to be the most critical for children with ASD?
- Would an SG, based on artistic representations and user-centered design, confronting ASD children with sound hypersensitivity to unpleasant auditory stimuli, be effective in reducing their sensitivity to sounds?
- To what extend this adventure game is appropriate for that specific purpose?

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 54–63, 2017.
https://doi.org/10.1007/978-3-319-71940-5_5

2 Related Works

We have categorized this analysis of related work into two sections exploring the goals and effectiveness of SGs that have been developed for autistic children. This analysis relates to studies concerning efficiency of SGs in moderating sensory hypersensitivity in ASD children.

2.1 Purpose of Serious Games for Children with ASD

In the past decade, the development of SGs for children with ASD has gradually increased [5]. A large number of studies have demonstrated significant improvements in the social interaction, communication and conversation skills of children with ASD after playing SGs [6–9]. For instance, PixTalk [8] is a smart phone application, with many pictures designed to help children express their feelings and emotions by browsing and selecting images. Likewise, ECHOES [9] is an intelligent application, in which autistic children interact with an avatar verbally, non-verbally, or both via a multi-touch display.

SGs for autistic children have two main purposes: education and therapy [11, 12]. According to our previous review of SGs for ASD children [10], there is a need for more research on therapy [11, 12]. One example targets autistic children that speak very quickly, who are likely to be unintelligible, by developing their speech to become clearer and more understandable [11]. Another SG combines tangible and graphical user interfaces on a PC, aimed at enhancing children's social and cognitive skills [12]. Although these SGs show significant promise for ASD therapy, further research is required.

2.2 Effectiveness of Serious Games in Moderating Sensory Hypersensitivity in Children with ASD

Children with ASD often struggle with Moderating Sensory Hypersensitivity (MSH) in one or multiple sensory modalities, including tactile, visual, vestibular and auditory modalities [13, 14]. Despite increased research on ASD, there is still a lack of understanding concerning the functioning of SGs when used in the context of ASD with MSH [10]. Nevertheless, to the best of our knowledge, there is only one published study in this area, focusing on vision [15]. This study used virtual reality as a sensory integration therapy (SIT), studying the relationship between sensory experiences and motor and behavioral performance. The study considers social skills training, visuomotor coordination ability assessment, and SIT. The game measures "visuomotor coordination ability"; for example, the player must break out of a virtual balloon using a tangible (real) stick. Additionally, the study has another session also created for SIT purposes. The application design was based on photos and recorded videos of an amusement park and playground. The game (EASe Funhouse Treasure Hunt[1]) uses 3D game technology to help ASD children and all other children diagnosed with auditory

---

[1] www.vision-play.com/product/ease-funhouse-tresure-hunt.

hypersensitivity, and sensory integration disorder cope with noise and improve sensory processing. Several studies have suggested the effectiveness of therapeutic games for children with ASD [16, 17], although more research is required.

3 Our Work

Although many autistic children are thought to suffer from auditory hypersensitivity [18] according to our present knowledge, no studies have discussed what the most relevant noises are, since there are large individual differences [19]. Accordingly, our research is categorized into two experiments: (1) Preliminary research which aims at determining the major characteristics of the sounds that affect sound sensitivity of children with ASD, and (2) A study which implements those sounds into a custom-designed SG, in order to explore the effectiveness of this game to familiarize and desensitize autistic children with sound hypersensitivity to critical sound cues.

3.1 Preliminary Research: Identify the Most Critical Sounds for Children with ASD

Ten male children already diagnosed with ASD, from small specialized autism organizations in Glasgow, UK, were asked to rate their perception of specific sounds using a Likert scale. Based on literature review and expert opinion (authors MP and DS), fifteen types of sounds were selected from [20], representing the most disturbing sounds according to people with auditory hypersensitivity, and considered as two independent variables; low-frequency sounds (drilling, traffic noise, dog barking) and high-frequency sounds (rattling of dishes, child crying, rustling of paper, applause, birdsong). Most of the audio resources were collected from the Simulation and Visualization library of the Glasgow School of Art, while a few audio samples were used from the Freesound website[2]. In terms of ensuring auditor level safety, we used Reaper 2016 software[3] to measure all 30 sound levels to be no more than 85 dB (normal hearing) [21], whilst Adobe Audition CC 2016 was utilized for editing, such as duplicating some audio samples (doors and school bells). A series of emoji were designed through Adobe Illustrator CS6 and Photoshop CS6, to translate the Likert scale items into graphical information that children could easily understand. We developed an Audio Interactive Questionnaire (AIQ) for Mac/Windows computers via Unity 2016[4] software (Fig. 1). The order of the sounds was coded and randomized, in order to avoid any possible order effect.

The AIQ data is presented in Table 1. The auditory questionnaire determines the independent variables to be considered in the primary experiment. As indicated in the following table, seven sounds have been highlighted as those being most critical.

[2] www.freesound.org.

[3] www.reaper.fm.

[4] www.unity3d.com.

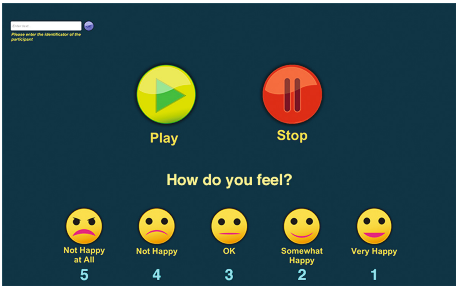


Fig. 1. Audio Interactive Questionnaire

Most critical sounds were:

1. Baby crying 2: sound of a new-born baby crying
2. Bell 2: sound of school bell ringing twice
3. Drilling 1: sound of drilling teeth
4. Drilling 2: sound of drilling wood
5. Fire engine 2: sound of three fire engine sirens
6. Traffic noise 1: sound of three cars passing on a highway
7. Train 1: sound of a moving train

These sounds were implemented in the SG described in the next sections.

3.2 Serious Game Study: Research Methods

An SG, Sinbad and the Magic Cure, was developed for Android devices. Prior to experimentation, the game was pilot-tested on typical children (n = 28) of the same age range as our target audience (8–11) at the Glasgow Science Center for three days (Fig. 2). The goal of this test was to determine the appropriate experimental duration, explore any difficulties or issues while playing the game on a tablet, assess interface accessibility by exploring the usability of touched buttons, and test the sound volume. Overall, the pilot was successful, with no technical issues or complications. Eventually, children were handed star stickers as a reward for their participation. Pilot outcomes were:

1. The game must be divided into three difficulty modes: easy, medium and difficult.
2. Control buttons need to be bigger, as they were too small for adequate control.
3. Sound volume was too low.

3.2.1 Participants

The main study was performed at a local primary school for children with special needs in Glasgow, UK. A new cohort of participants was built using seven children aged 8–11 (6 males, 1 female) with ASD and hypersensitivity to sounds but noun of them had sensitivity to visual stimuli, that were involved in our preliminary research (Sect. 3.1). The main researcher, along with the gate-keepers (teachers and teaching assistants from the school), observed the children throughout the experiment. Due to space shortage, participants were divided into two groups of three and four. Both groups were exposed successively to the same experimental treatment.

Table 1. Most critical Sounds for ASD children aged 8–11 years old

Sounds	Mean Rating (1-5)	σ
Applause 1	2.2	1.14
Applause 2	2.4	0.7
Baby Crying 1	3.2	1.32
Baby Crying 2	3.7	1.64
Bell 1	2.8	0.79
Bell 2	3.4	0.84
Bird 1	1.4	0.7
Bird 2	2.3	1.16
Checkout Till 1	3	0.94
Checkout Till 2	2.3	1.06
Dog 1	2.2	1.32
Dog 2	2.4	1.07
Door slamming 1	2.6	1.07
Door slamming 2	2.6	1.07
Drilling 1	3.5	0.97
Drilling 2	3.4	1.58
Fire engine 1	3.2	1.32
Fire engine 2	4.1	1.29
Fireworks 1	3	1.63
Fireworks 2	2.2	1.75
Hammering 1	3.1	1.29
Hammering 2	2.7	1.34
Rattling of dishes 1	2.8	1.03
Rattling of dishes 2	2.8	1.4
Rustling paper 1	2	0.94
Rustling paper 2	2.5	0.97
Traffic noise 1	3.9	0.99
Traffic noise 2	2.5	1.72
Train 1	3.4	1.07
Train 2	2.6	1.65

3.2.2 Materials

The game's story and design are inspired by the Arabic fairy tale Sinbad the Sailor from 1001 Nights[5], and combine Islamic and contemporary art. The development

[5] www.wollamshram.ca/1001.

Fig. 2. Pilot-test at Glasgow Science Centre, 17-19/06/2017

mixes artistic and game-engine platforms such as SketchBookPro 2016[6], Photoshop CS6, Illustrator CS6 and Unity 3D 2016 software. The participants were required to play the game on a Fire tablet, 7.0 in. display, with headphones (Fig. 3).

Fig. 3. Sinbad and the Magic Cure, a SG for autistic children with audio hypersensitivity

3.2.3 Procedure

Children were required to play with the SG throughout multiple sessions. Having previously obtained formal consent from their caregivers, prior to each session, the purpose of the study was explained to the children and they were informed about their rights as participants to withdraw from the session at any point. Participants were given the opportunity to ask questions or raise any concerns throughout the study. In the first session, children were asked to fill out the AIQ in the same way as for session 3.1 on a laptop for about 10 min, and they were then given tablets and headphones. Also, the

[6] www.autodesk.com/.

aim of playing the game, and how to play it, was explained to participants. The tasks during each session were:

1. Mood table: children were asked how they felt prior to playing the game, in order to fill out the mood table in (Fig. 4); this step is important as it could affect participants' answers.

Date	Duration of playing the game	How do you feel?					Rewards
		Not Happy at All	Not Happy	Natural	Somewhat Happy	Very Happy	

Fig. 4. Mood table

2. Gameplay: Playing our Sinbad and the Magic Cure game for approximately 25 min. The game involves six levels, each having a different number of sounds (minimum of two and maximum of four sounds per level). The aim of the game is for the avatar (Sinbad) to collect various sound objects in order to exchange them with a wicked witch for a magic cure – as a result, his friend Yasmin turns back into a girl again. The objects appear for a few seconds before vanishing and emerging again in new locations in the digital environment of the game (Fig. 5).

Fig. 5. Group 1 during playing the game Sinbad and the Magic Cure

3. AIQ: At the end of every session, participants rated the 30 sounds based on their feeling in the same way as for session 3.1 on a laptop for about 10 min.

Every group took around 30 min in total to complete playing the game and filling out the AIQ. Finally, the children were thanked for their contribution and given a small gift (star stickers) as appreciation.

3.3 Data Analysis

A typical Likert scale (Very Happy = 1, Somewhat Happy = 2, OK = 3, Somewhat Unhappy = 4 and Not Happy at All = 5) using iconic faces as in the preliminary research (Sect. 3.1), was employed to rate the statements proposed through the questionnaires.

4 Results

Generally, we observed that participants enjoyed playing the game very much. They tended to keep playing even after the required time. Additionally, they were keen to have the game installed on their mobile devices in order to play it at home; for example, they asked when the game would be published in Google Play so that they could download it to their own devices. Indeed, they were looking forward to having more levels for the game after they finished, and kept asking whether there was a new level available for playing. They engaged well with the game during the duration of the experiment. Interestingly, although DM mentioned that he preferred to play 3D games, he nevertheless kept playing our (2D) game. Participants were competitive and enthusiastic to finish the game. Six out of seven were able to do this in approximately 20 min, and only one participant could not finish the game.

The AIQ mean and standard deviation data are presented in the Table 2. All participants, as mentioned earlier, voluntarily played the game for seven days; however, one participant (with co-morbid Down Syndrome) was excluded from data analysis since he was happy with all seven sounds and found the game too difficult to complete in the easy mode, although he never actually stopped playing. The sound tolerance of these six participants was measured on a day-to-day basis after finishing playing the SG (Table 2). Overall, the results support the suggestion that playing SGs over a long period of time can be an effective way for audio hypersensitive autistic children to develop

Table 2. Mean and Standard Deviation of the critical sounds

Critical Sounds		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Baby Crying 2	M	4.000	4.167	3.500	3.500	3.000	4.000	3.500
	σ	0.632	0.983	1.975	1.225	0.632	1.265	1.378
Bell 2	M	3.167	3.333	2.667	2.500	3.333	3.000	2.667
	σ	1.472	1.211	1.633	1.049	1.506	1.673	1.366
Drill 1	M	3.833	3.667	3.833	3.667	3.167	3.167	3.000
	σ	0.753	0.816	1.169	0.816	0.408	0.753	0.632
Drill 2	M	3.333	3.333	3.000	3.667	3.000	3.167	3.167
	σ	1.033	1.095	1.095	1.033	1.265	1.602	0.983
Fire Engine 2	M	4.333	3.833	4.333	3.667	3.667	4.333	4.167
	σ	0.816	1.169	1.033	0.816	1.211	0.816	0.983
Traffic Noise 1	M	3.333	2.833	3.333	2.333	3.167	3.000	3.667
	σ	1.211	0.753	0.816	0.816	0.983	0.894	0.816
Train 2	M	2.833	2.833	3.667	2.333	2.333	2.833	2.833
	σ	1.602	0.983	1.211	1.033	1.033	1.602	0.983

tolerance to sounds. Interestingly, ratings for the sounds of baby crying 2, bell 2, drill 1, drill 2, and fire engine 2 decreased. In addition, traffic noise 1 statistics were unstable in contrast with train 2 ratings (Table 2). Although, none of these changes suggested a drastic variation of trend in the context of this experiment, these results remain nonetheless promising for further experimentation.

5 Discussion and Future Work

There is much evidence to suggest that serious games provide an effective intervention therapy tool. However, more studies utilizing SGs for ASD children with moderating sensory hypersensitivity are required. Although further research is needed to explore the eventual effect of order and increase exposure to the game, this study provides evidence that using an SG in the longer term could improve sound tolerance in children with ASD. The results confirm that mobile devices may provide a suitable supportive tool for children suffering from ASD and sound hypersensitivity. Participating children reported feeling more accustomed to the sounds by the end of the experimental week. Moreover, they showed interest and engagement with the game. Indeed, as most children are accustomed to playing games on mobile devices, they become familiar with such games very quickly. One limitation of this study is the small sample size and the short duration of the study. In the future, we will repeat the experiment with larger groups for longer periods of time. We will also further develop the game, adding more levels to maximize gameplay duration. The game will also be customizable, and based on children's needs, so that parents can select the most unpleasant sounds to implement in the game, in order to increase its effectiveness.

Acknowledgments. This research was fully funded by The King Abdullah Scholarship Program (2014–2017). We would like to thank the children, parents, and teachers for participating in this research. We would also like to thank June Grindley, and Neelam Jhakra, Dr. Sandy Louchart, and Brian Loranger for their feedback and suggestions. In particular, we would like to thank the depute head teacher and other staff from the participating school, Creatovators CIC, Strathclyde Autistic Society Club, Glasgow Science Center, and the School of Simulation and Visualisation of the Glasgow School of Art for all their help in conducting this research. Special thanks for Victor Portela Romero, Louise Dolan, Hajara Alfa, Akash Angral, Dr. Polina Zioga, Dr. Jessica Argo, and Sawitree Wisetchat for their invaluable supports.

References

1. American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders. Arlington (2013). https://doi.org/10.1176/appi.books.9780890425596.744053
2. CDC: CDC estimates 1 in 68 children has been identified with autism spectrum disorder. CDC Newsroom (2014)
3. NHS Autism spectrum disorder - NHS Choices, http://www.nhs.uk/conditions/Autistic-spectrum-disorder/Pages/Introduction.aspx. Accessed 22 July 2017
4. Gomes, E., Pedroso, F.S., Wagner, M.B.: Auditory hypersensitivity in the autistic spectrum disorder. Pró-Fono Rev Atualização Científica **20**, 279–284 (2008). https://doi.org/10.1590/S0104-56872008000400013

5. Boucenna, S., Narzisi, A., Tilmont, E., et al.: Interactive technologies for autistic children: a review. Cognit. Comput. **6**, 722–740 (2014). https://doi.org/10.1007/s12559-014-9276-x
6. Hourcade, J.P., Williams, S.R., Miller, E.A., et al.: Evaluation of tablet apps to encourage social interaction in children with autism spectrum disorders. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2013, pp. 3197–3206 (2013). https://doi.org/10.1145/2470654.2466438
7. Gal, E., Lamash, L., Bauminger-Zviely, N., et al.: Using multitouch collaboration technology to enhance social interaction of children with high-functioning autism. Phys. Occup. Ther. Pediatr. **36**, 46–58 (2016). https://doi.org/10.3109/01942638.2015.1040572
8. De Leo, G., Gonzales, C.H., Battagiri, P., Leroy, G.: A smart-phone application and a companion website for the improvement of the communication skills of children with autism: clinical rationale, technical development and preliminary results. J. Med. Syst. **35**, 703–711 (2011). https://doi.org/10.1007/s10916-009-9407-1
9. Bernardini, S., Porayska-Pomsta, K., Smith, T.J.: ECHOES: An intelligent serious game for fostering social communication in children with autism. Inf. Sci. (Ny) **264**, 41–60 (2014). https://doi.org/10.1016/j.ins.2013.10.027
10. Zakari, H., Ma, M., Simmons, D.: A review of serious games for children with Autism Spectrum Disorders (ASD). Serious Games Dev. Appl. SE - **9**(8778), 93–106 (2014). https://doi.org/10.1007/978-3-319-11623-5_9
11. Hoque, M., Lane, J., Kaliouby, R., et al.: Exploring speech therapy games with children on the autism spectrum, pp. 1–4 (2014)
12. Barajas, A.O., Al Osman, H., Shirmohammadi, S.: A serious game for children with autism spectrum disorder as a tool for play therapy. https://doi.org/10.1109/SeGAH.2017.7939266
13. Tomchek, S., Dunn, W.: Sensory processing in children with and without autism: a comparative study using the short sensory profile. Am. J. Occup. **61**, 190–200 (2007)
14. Case-Smith, J., Weaver, L.L., Fristad, M.A.: A systematic review of sensory processing interventions for children with autism spectrum disorders. Autism **19**, 133–148 (2014). https://doi.org/10.1177/1362361313517762
15. Choi, M., Lim, C.: Interactive therapy system design for children with autistic. Therapy (2010)
16. Rogers, S.J., Ozonoff, S.: Annotation: what do we know about sensory dysfunction in autism? A critical review of the empirical evidence. J. Child Psychol. Psychiatry **46**, 1255–1268 (2005). https://doi.org/10.1111/j.1469-7610.2005.01431.x
17. Dawson, G., Watling, R.: Interventions to facilitate auditory, visual, and motor integration in autism: a review of the evidence. J. Autism Dev. Disord. **30**, 415–421 (2000). https://doi.org/10.1023/A:1005547422749
18. Gomes, E., Pedroso, F.S., Wagner, M.B.: Auditory hypersensitivity in the autistic spectrum disorder. Pró-Fono Rev Atualização Científica **20**, 279–284 (2008). https://doi.org/10.1590/S0104-56872008000400013
19. Ben-Sasson, A., Liat, A., Ae, H., et al.: A meta-analysis of sensory modulation symptoms in individuals with autism spectrum disorders. https://doi.org/10.1007/s10803-008-0593-3
20. Anari, M., Axelsson, A., Eliasson, A., Magnusson, L.: Hypersensitivity to sound Questionnaire data, audiometry and classification. Scand. Audiol. **28**, 219–230 (1999). https://doi.org/10.1080/010503999424653
21. Headphones and the Noise at Work Regulations, LexisNexis Butterworths (1992). https://www.isvr.co.uk/reprints/Headphones%20and%20the%20nawr%20p.pdf. Accessed 14 Feb 2017

Identifying Different Persuasive Gaming Approaches for Cancer Patients

Teresa de la Hera Conde-Pumpido(✉)

Erasmus University Rotterdam, Woudestein, Van der Goot Building, M8.35, P.O. Box 1738, 3000 DR Rotterdam, Netherlands
delahera@eshcc.eur.nl

Abstract. Using as a starting point behavior scientist B.J. Fogg's conceptual framework on the role computer technology plays for users, in this paper it will be argued that persuasion through digital games can be approached from three different perspectives: digital games can be used as media for persuasion, digital games can be used as tools for persuasion and digital games can be used as social actors for persuasion. In this paper, I use four cancer gaming cases to illustrate how these three different persuasive roles can be used to accomplish different persuasive goals. The categorization proposed in this paper can serve to clarify what we mean when we refer to persuasion in relation to digital games and can be used as analytical approach for the study of persuasive gaming strategies.

1 Introduction

The persuasive potential of digital games has been proven to be useful to change, reinforce or shape the attitude and/or behavior of players in several fields such as advertising [1, 2], pro-social communication [3] or healthcare [4]. However, if we pay attention to the different academic definitions used for the concept of persuasive games, and the different categories of persuasive goals studied within this field, it can be concluded that researchers mean different things when they refer to the persuasive potential of digital games.

Persuasive games have been defined as "games that mount procedural rhetoric effectively" [5], "games explicitly created to change attitudes and behavior" [3], "digital games that aim to shape, reinforce or change the perceptions, emotions, beliefs, behavioral intentions and behaviors of players" [6], "games that aim to increase players' awareness of critical and timely social issue" [7], or "designed interventions with the primary purpose of changing a user's behavior or attitude in an intended way" [4]. It follows that, while some researchers relate persuasion through digital games to their capacity to convey persuasive messages [8], others focus their attention on their

This work is part of the research programme "Persuasive Gaming in Context: From theory-based design to validation and back" with project number 314-99-106 which is financed by the Netherlands Organisation for Scientific Research (NWO).

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 64–72, 2017.
https://doi.org/10.1007/978-3-319-71940-5_6

capacity to trigger specific behaviors [4] or facilitating specific interactions among players [9].

These differences in the way persuasion through digital games is defined and studied are the result of the complex nature of this practice. This complexity is not only due to the wide range of possible applications but also due to the complexity of the process of persuasion itself and how the specificities of digital games have an influence in this process. For this reason, it is necessary to provide theoretical frameworks that can serve not only to analyze the use of different persuasive strategies in relation to different persuasive goals but also that can become a useful tool to select a specific persuasive approach when designing persuasive games. The provision of new theoretical concepts and frameworks is of special relevance in the relatively new field of persuasive games, as these become relevant tools for scholars in the field, that can serve to better structure their work and communicate about their insights. This is, therefore, the main goal of this paper.

In this paper I use as a starting point the conceptual framework of behavior scientist B.J. Fogg [10], who described the overlap between persuasion and interactive technology, to argue that it is possible to identify three different roles of digital games when used with persuasive intentions. Digital games can be used as media for persuasion providing compelling meaningful experiences that convey specific messages, digital games can be used as tools for persuasion to persuade players in a number of ways such as making a target behavior easier to perform, and digital games can be used as social actors for persuasion "by applying the same persuasion principles that humans use to influence others" [10]. Fogg's framework, that is based on the role computer technology plays for users, does not explain how specific persuasive potentials of digital games can be linked to specific persuasive roles, a question that I address in this paper.

In order to illustrate my arguments, I will analyze how these three roles of digital games have been used in the field of healthcare, and specifically for the design of cancer games. In this paper I will use the *Re-Mission Game* [11], *The Cancer Game*, the *Cogmed Memory Training* [12] and *The Survivor Games* [13] as examples of how this topic can be addressed from different approaches when different persuasive roles of digital games are used to promote desired *sick-role behaviors* in cancer patients.

2 Digital Games to Support Cancer Patients

The positive effects of the use of video games in the field of health is a reality demonstrated by numerous research studies worldwide [14]. Video games are being used with excellent results to, among other things, overcome phobias [15], support rehabilitation therapies [16] or carry out neuropsychology therapies [17]. Among all its applications in the field of health, it stands out its use to help cancer patients (especially children and adolescent) to adhere to its treatment.

According to the World Health Organization cancer is among one of the leading causes of mortality worldwide, with 14 million new patients in 2012, a figure that is expected to increase by about 70% over the next two decades. There is extensive knowledge about the causes of cancer, how it can be prevented and how to manage the disease once it has been detected. Healthcare persuasion uses this knowledge to promote *sick-role behaviors*. *Sick-role behaviors* are those undertaken by cancer patients to succeed in their recovery or to improve their quality of life while under treatment [18]. Healthcare persuasion to promote sick-role behaviors in cancer patients includes strategies to help patients to adhere to their treatments (e.g. taking a full course of oral chemotherapy), and to manage their disease (e.g. controlling anxiety, reducing stress, taking a balanced diet) [4].

Previous research has shown that digital games are an effective vehicle for cancer-related healthcare persuasive strategies [14]. Cancer treatments, especially chemo, are really aggressive treatments and with many side effects not always easy to cope with. For this reason, and especially in the case of children, support adherence to treatment becomes especially relevant. In this respect, there are three different challenges to overcome, and related to each of these challenges, there is a role that video games can fulfill to complement the support to the adhesion. The three challenges to which I refer are: lack of information, lack of motivation and difficulties in coping with treatment. In the following sections, I argue how digital games can play three different roles to support cancer patients to face these three challenges.

3 Digital Games as Media for Persuasion

The first challenge that digital games can help overcome is the lack of information about treatment, i.e. how it works and what effects it may have. There are numerous studies that show that many cancer patients have a misperception about their effects before starting treatment [19], often fearing worse side effects than those who are actually going to experience. For this reason, it has been demonstrated that providing complete and accurate information to the patient's reality is essential for adherence to treatment [20]. When we talk about children or adolescents, overcoming this challenge is more complicated than in the case of adults [20]. That is why looking for creative ways to offer treatment information, which is understandable to them and does not cause them fear, is something that especially worries health specialists. Digital games have proved to be an interesting media to transmit the desired information to children in a way that is well received by them.

In their role as media for persuasion, digital games can be used to promote *sick-role behaviors* of cancer patients by providing compelling experiences that help them to better understand what are they going through and how the therapies and life styles they are asked to stick to, can help them to overcome their disease. According to Fogg's [10] theoretical framework, this can be done by (1) allowing people to explore cause-effect relationships, (2) by providing people with vicarious experiences that motivate or (3) helping people rehearse a behavior.

An example on how digital games has been used already for this purpose, is the *Re-Mission Game* (see Fig. 1), designed to increase knowledge about chemotherapy, change perceptions of patient's ability to influence health outcomes, and to provide patients with confidence in their ability to meet the specific demands of cancer treatment and recovery. In the game, children control a character that navigates through the interior of the human body. Through the character, they can destroy cancer cells and

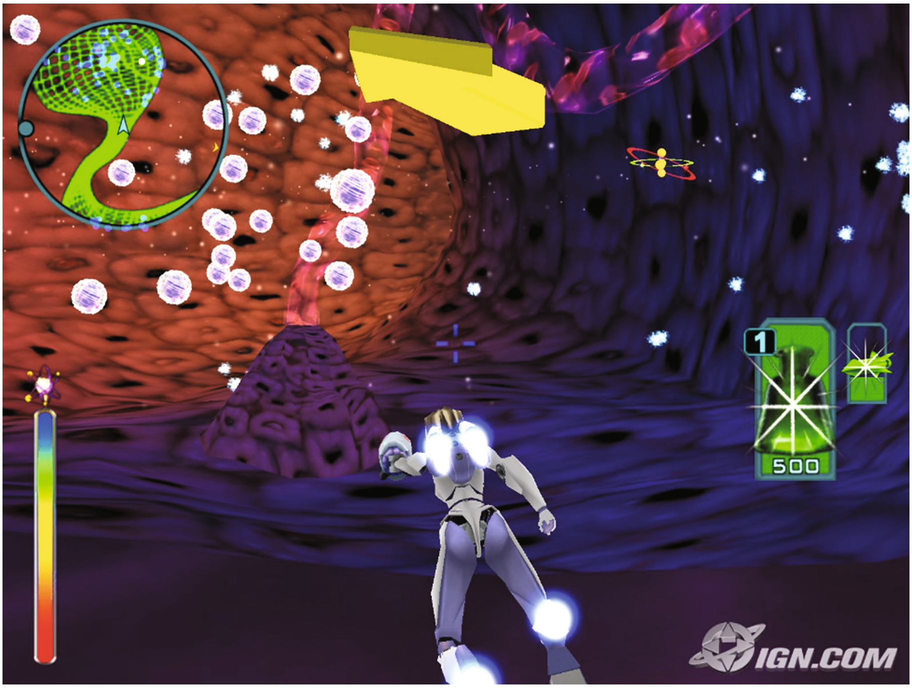


Fig. 1. Screenshot *Re-Mission*

defeat the main tumor. Through friendly design and simple game mechanics, children can understand a complex process.

This game becomes this way a medium for persuasion by allowing children to explore the cause-effect relationships of chemotherapy treatments. The mechanics of the game and the challenges they need to complete, allow children to explore through their actions in the game, which are the effects of chemotherapy in their bodies. By better understanding how this process work, they can be more open to undergoing their treatment, regardless of the side effects the treatment has in their bodies. Research on this game has indeed shown that *Re-Mission* is effective in increasing adherence to treatment among its players [21].

4 Digital Games as Tools for Persuasion

The second challenge is related to the lack of motivation to start or continue a treatment. As I mentioned earlier, cancer treatments or their side effects are often aggressive or difficult to cope with. The diagnosis and treatment of cancer is a stressful and threatening experience for children [22]. Although survival rates for childhood cancer are higher than ever before, the course of treatment for cancer, such as chemotherapy, surgery or radiotherapy, is still a very stressful experience in the life of a child. However, cancer patients may not only be at risk from adverse medical effects, their psychosocial well-being may also be severely affected as a result of cancer and its therapy [23]. For this reason, even if patients have enough information on how a treatment works and its benefits, they may not have enough motivation to start or continue at a given time. Digital games can help in this regard, serving as a tool to motivate patients to comply with treatment guidelines.

During the past decade, there has been an increase in the use of therapeutic play intervention to help cancer patients cope with the stress of hospitalization and treatment [24]. Therapeutic play is a set of structured activities designed according to psychosocial and cognitive development of cancer patients and health-related issues to help them to cope with psychological and emotional difficulties [25]. The central goal of therapeutic play is to facilitate the emotional and physical well-being of patients [26].

In their role as tools for persuasion, digital games are designed to influence and motivate people in specific ways by making activities easier or more efficient to do [10]. Patients are often required to undergo procedures or engage in behaviors that are painful and aversive on the one hand (e.g., undergoing chemotherapy) or boring and mundane on the other (e.g., taking pills, exercising on a regular basis). These procedures and behaviors are often necessary to maintain and improve health or even to cure the patient's disease [14, p. 113].

According to Fogg's theoretical framework [10] in their role as tools for persuasion, digital games can be used to support cancer patients in undergoing painful or boring procedures by (1) reducing a complex activity to a few simple steps, (2) leading users through a predetermined sequence of actions or events, step by step, (3) suggesting a behavior at the most opportune moment, (4) helping people to monitor themselves to modify their attitudes or behaviors to achieve a predetermined goal or outcome, or (5) allowing doctors to monitor the behavior of patients.

An example of this is the game *Cogmed Memory Training* (see Fig. 2), developed to aid in the adherence to treatment of patients with cognitive problems, one of the side effects of treatments such as brain chemotherapy, among others. Neuropsychological rehabilitation treatments require the patient to repeat a series of simple tasks (such as associating objects of the same color and shape), in a continuous and intense way. The treatment can be really exhausting for patients, and it can become boring for children. For this reason, neuropsychologists are using games that help to work the same skills in a more entertaining way for patients. The game *Cogmed Memory Training* has shown, in this regard, to be an effective tool to support adherence to the treatment.

Going back to Foog's theoretical framework, the game helps to lead patients through the different steps of the treatment, this is done trough a series of mini-games focused on training different skills in patients, necessary for their recovery. The game keeps patients busy and focused on different challenges and levels, which help them to persist in the treatment. Besides this, the game helps patients and doctors to monitor their progress, which helps patients to increase their level of perception to be in control of their own health, and also helps doctors to adapt the treatment considering patients' progress.

Another example of how digital games can be used as tools for persuasion is *The Cancer Game* (Kristula & Oda, 2003), an online game designed to provide stress relief for cancer patients, in which they can visualize and destroy cancer on a computer

Fig. 2. Screenshot *Cogmed Memory Training*

screen. The game was designed by researcher Dave Kristula after having to deal with a cancer treatment himself. The game, developed in collaboration of Professor Yuko Oda, was designed based on previous research on the healing through visualization therapies focusing on stress relief.

The game fosters a specific behavior while the game is played: visualizing how cancer can be destroyed and helping patients to envision that desired moment. Besides this, the game leads also patients through a simple but predetermined sequence of actions, that help them to keep focus their attention on a simple activity and leave aside other thoughts that might increase anxiety.

5 Digital Games as Social Actors for Persuasion

The third challenge to which I referred at the beginning of the article was the difficulty in coping with treatment. It may be the case of a patient who has all the information about the treatment and who also has the motivation to carry it out, but who is in difficulties in the moment of coping. Examples of these difficulties may be anxiety or fear of how the disease will develop, discomfort generated by the treatment received or hair loss in the case of patients treated with chemotherapy. In this respect, the support of other patients who are going through or have experienced the same difficulties can be of particular help. However, it is not always easy to establish new relationships with other patients, or initiate conversations on issues that concern you with people with whom you do not have a close relationship. To support patients to face this challenge, digital games can take the role of social actors for persuasion and be used for the purpose of mediating the process of seeking social support. The persuasive potential of this role relies on the capacity of digital games to be used to persuade players "by giving a variety of social cues that elicit social responses" from them [10].

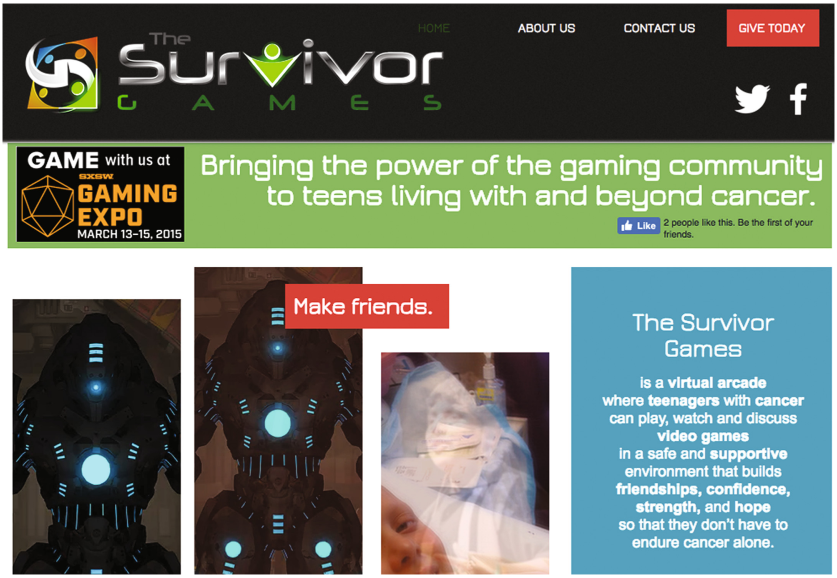


Fig. 3. Screenshot The Survivor Games

When used as social actors for persuasion, digital games may utilize either computer-human persuasion or computer-mediated persuasion [27]. Although computers cannot communicate in the same way as humans, there are studies that suggest that computer-human persuasion may utilize some patterns of interaction similar to social communication [28], whereas computer-mediated persuasion means that people are persuading others through computers, e.g. discussion forums, e-mail, instant messages, blogs, or social network systems. Recent research, for example, shows that dialogue support features play a significant role in relation to adherence to web-based health interventions, especially for young adults, who tend to be influenced by their peers [9].

An example of how digital games can play the role of social actors for persuasion is the platform *The Survivor Games* (see Fig. 3). This platform was created as a community aimed at patients and survivors of childhood cancer. On the web, its users can play a selection of multiplayer video games specially chosen for them. Video games become the point of union between players, who begin to talk about their games and the characters of their games, and on many occasions that the first contact evolves until a real friendship is established between two people who are going through similar difficulties. Thus finding a support for coping with them. This is related to the capacity of digital games to foster social interaction [29], and the tools implemented in the platform itself to facilitate this interaction, such as the chat box that allows players to chat while playing. Besides this, the context in which these games are played, a platform especially aimed at young cancer patients and cancer survivors, also becomes relevant in this process, as users might find it as a natural context to look for contact with other people going through similar experiences, which can help them to feel more comfortable to disclose their personal experiences, share emotions and feelings and look for support. We can talk then about game-mediated type of persuasion, meaning that the game becomes persuasive when being played in a specific context with a specific purpose [30].

6 Conclusions

The four games mentioned in this article are just three examples of how digital games can help overcome the three challenges related to adherence to cancer treatments. Through these cases, I have illustrated how digital games can play three different roles in the process of persuasion, becoming media, tools and social actors for persuasion. The theoretical framework presented in this paper can serve to better understand how these roles can serve to achieve different goals when games are used for persuasive purposes. In a future paper, I will explain how specific properties of digital games are linked to each of these three roles.

References

1. Deal, D.: The ability of branded online games to build brand equity: an exploratory study. In: DiGRA 2005 Conference: Changing Views-Worlds in Play (2005)
2. Wise, K., Bolls, P.D., Kim, H., Venkataram, A., Meyer, R.: Enjoyment of advergames and brand attitudes: the impact of thematic relevance. J. Interact. Advert. **9**(1), 27–36 (2008)
3. Ruggiero, D.: The effect of a persuasive game on attitude towards the homeless. Children (2014)
4. Orji, R., Mandryk, R.L., Vassileva, J., Gerling, K.M.: Tailoring persuasive health games to gamer type. In: Proceedings SIGCHI Conference on Human Factors in Computing Systems - CHI 2013, pp. 2467–2476 (2013)
5. Bogost, I.: Persuasive games on mobile devices. In: Fogg, B.J., Eckles, D. (eds.) Mobile Persuasion, Stanford: Stanford University (2007)
6. De la Hera Conde-Pumpido, T.: Persuasive structures in advergames: conveying advertising messages through digital games, Utrecht University (2014)
7. Kaufman, G., Flanagan, M., Seidman, M.: Creating stealth game interventions for attitude and behavior change: an 'Embedded Design' model. In: Proceedings DiGRA 2015 Divers. Diversity of play: Games – Cultures – Identities, pp. 1–13 (2015)
8. Bogost, I.: Persuasive Games: The Expressive Power of Videogames. MIT, Cambridge (2007)
9. Kulyk, O., Den Daas, C., David, S., Van Gemert-Pijnen, L.: How persuasive are serious games, social media and mHealth technologies for vulnerable young adults? Design factors for health behavior and lifestyle change support: Sexual health case. In: CEUR Workshop Proceedings, vol. 1369, pp. 28–42 (2015)
10. Fogg, B.J.: Persuasive technology: using computers to change what we think and do. In: Persuasive Technology: Using Computers to Change What We Think and Do, vol. 5, no. 1, p. 283 (2003)
11. Hope Lab: Re-Mission (2006)
12. Pearson: Cogmed Memory Training
13. The Survivor Games

14. Kato, P.M.: Video games in health care: Closing the gap. Rev. Gen. Psychol. **14**(2), 113–121 (2010)
15. Miloff, A., Lindner, P., Hamilton, W., Reuterskiöld, L., Andersson, G., Carlbring, P.: Single-session gamified virtual reality exposure therapy for spider phobia vs. traditional exposure therapy: study protocol for a randomized controlled non-inferiority trial. Trials **17** (1), 60 (2016)
16. Reid, D.T.: Benefits of a virtual play rehabilitation environment for children with cerebral palsy on perceptions of self-efficacy: a pilot study. Pediatr. Rehabil. **5**(3), 141–148 (2002)
17. Hardy, K., Bonner, M., Willard, V.: Computerized cognitive training for survivors of pediatric cancer. Pediatr. Blood Cancer **55**(5), 776–777 (2010)
18. Baranowski, T., Blumberg, F., Buday, R., DeSmet, A., Fiellin, L.E., Green, C.S., Kato, P. M., Lu, A.S., Maloney, A.E., Mellecker, R., Morrill, B.A., Peng, W., Shegog, R., Simons, M., Staiano, A.E., Thompson, D., Young, K.: Games for health for children-current status and needed research. Games Heal. J. **5**(1), 1–12 (2015)
19. Wakefield, C.E., Butow, P., Fleming, C.A.K., Daniel, G., Cohn, R.J.: Family information needs at childhood cancer treatment completion. Pediatr. Blood Cancer **58**(4), 621–626 (2012)
20. Mitchell, W., Clarke, S., Sloper, P.: Care and support needs of children and young people with cancer and their parents. Psycho-Oncology **15**(9), 805–816 (2006)
21. Kurt, A.S., Savaser, S.: An effect of re-mission video game on perceived stress levels of adolescents with cancer. Acta Paediatr. Int. J. Paediatr. **100**, 98–99 (2011)
22. Hicks, M.D., Lavender, R.: Psychosocial practice trends in pediatric oncology. J. Pediatr. Oncol. Nurs. **18**(4), 143–153 (2001)
23. Langeveld, N.E., Grootenhuis, M.A., Voûte, P.A., de Haan, R.J.: Posttraumatic stress symptoms in adult survivors of childhood cancer. Pediatr. Blood Cancer **42**, 604–610 (2004)
24. Li, W.H.C., Chung, J.O.K., Ho, E.K.Y.: The effectiveness of therapeutic play, using virtual reality computer games, in promoting the psychological well-being of children hospitalised with cancer. J. Clin. Nurs. **20**(15–16), 2135–2143 (2011)
25. LeVieux-Anglin, L., Sawyer, E.H.: Incorporating play interventions into nursing care. Pediatr. Nurs. **19**(5), 459–463 (1993)
26. Vessey, J.A., Mahon, M.M.: Therapeutic play and the hospitalized child. J. Pediatr. Nurs. **5** (5), 328–333 (1990)
27. Oinas-Kukkonen, H., Harjumaa, M.: Towards deeper understanding of persuasion in software and information systems. In: Proceedings of the 1st International Conference on Advances in Computer-Human Interaction, ACHI 2008, pp. 200–205 (2008)
28. Fogg, B.J., Nass, C.: How users reciprocate to computers : an experiment that demonstrates behavior change. In: CHI 97, pp. 331–332, March 1997
29. De la Hera Conde-Pumpido, T., Paz Aléncar, A.: Collaborative digital games as mediation tool to foster intercultural integration in primary Dutch schools. e-Learning Papers, vol. 43, pp. 13–23 (2015)
30. De la Hera Conde-Pumpido, T.: Persuasive gaming: identifying the different types of persuasion through games. Int. J. Serious Games **4**(1), 31–39 (2017)

Prosodiya – A Mobile Game for German Dyslexic Children

Heiko Holz[1,2(✉)], Katharina Brandelik[2], Jochen Brandelik[2], Benedikt Beuttler[2,3], Alexandra Kirsch[4], Jürgen Heller[3], and Detmar Meurers[1,5]

[1] LEAD Graduate School & Research Network, Eberhard Karls Unviersität Tübingen, Tübingen, Germany
{heiko.holz,detmar.meurers}@uni-tuebingen.de
[2] Tübinger Institut für Lerntherapie GmbH, Tübingen, Germany
{k.brandelik,j.brandelik}@til-lerntherapie.de
[3] Department of Psychology, Eberhard Karls Unviersität Tübingen, Tübingen, Germany
{benedikt.beuttler,juergen.heller}@uni-tuebingen.de
[4] Department of Computer Science, Eberhard Karls Unviersität Tübingen, Tübingen, Germany
alexandra.kirsch@hci.uni-tuebingen.de
[5] Department of Linguistics, Eberhard Karls Unviersität Tübingen, Tübingen, Germany

Abstract. Approximately 4–10% of the German population suffers from developmental dyslexia. The learning disorder afffects educational, personal, and social development of children in a negative way. Mobile serious games have the potential to support dyslexic primary-school children in addition to school support and learning therapy. We propose such a mobile serious game called "Prosodiya", with the help of which dyslexic children can improve their reading and writing performance. Prosodiya includes innovative and evidence-based interventions that focus on improving the awareness of linguistic features related to syllable stress. We report the results of a pilot study of a preliminary version of the game. Results indicate that the children enjoyed playing the game, that their motivation was maintained, and that they wanted more levels.

1 Introduction

Developmental dyslexia is one of the most frequent learning disorders [24]. Affected children suffer massively from an impaired literacy acquisition – compared to their classmates, they acquire reading and writing skills in a much slower pace and not as proficient [29]. Usually, they lose motivation for the learning process as well as faith that they will ever be able to develop a comprehension of literacy language [2]. If these children do not receive appropriate treatments, negative consequences in the long run may arise, such as poor graduation and

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 73–82, 2017.
https://doi.org/10.1007/978-3-319-71940-5_7

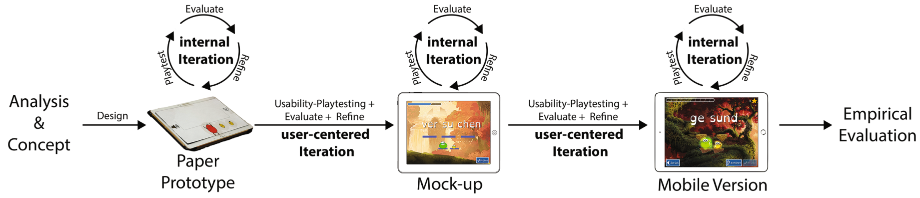


Fig. 1. Exemplary Iterative Children-Centered Game Design Process (ICCGD) for the game "stress pattern".

higher chances to drop out from school, resulting in poor employment prospects or unemployment [10]. Overall, the learning disorder has a negative impact not only on mental health but also on social and cultural participation [1]. Serious Games have the potential to address the aforementioned difficulties. In shaping an individual's learning curve, they boost motivation [11], and lead to successful learning processes (cf. [5]). With the great advantage of independence of time and location, mobile games can additionally help the children to overcome their learning disorder outside of learning therapy and classrooms – serious games for dyslexic children have been proven to have positive effects on the process of literacy acquisition (e.g., [4,18]).

In the first part of this article, we present such a mobile serious game for dyslexic children aged 6–12. Focus of this game called "Prosodiya" is on well-founded user-centered development, on the adaptivity of the digital interventions, and on embodied training. To the best of our knowledge, Prosodiya adds two novelties to the field of serious games for dyslexic children. First, it is the first digital therapy approach that focuses on improving the awareness of syllable stress and associates the stressed syllable's linguistic features to orthographic principles of the German orthography. Second, it is the first mobile game supporting embodied training using sensor-based gesture recognition.

In the second part of this article we present the results of a pilot testing a preliminary version of the game. Focus is on user experience and usability, but the impact on literacy development is also considered.

2 Prosodiya

To avoid failure caused by losing focus on the target audience, we particularly emphasized the involvement of primary-school children as the end user. We followed an approach called iterative children-centered game design (ICCGD, see Fig. 1) during the whole development process. The ICCGD combines the two familiar approaches of user-centered design [22] (UCD) and iterative game design [12] (IGD).

Prosodiya is based on recent empirical findings and on evidence-based interventions (e.g., [17,27]). For example, a main component is training phonological awareness as children with dyslexia often struggle with this basic skill [6,32].

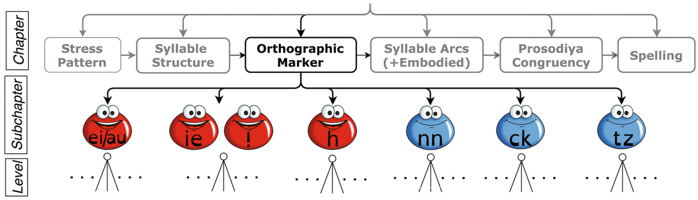


Fig. 2. Structure of Prosodiya. The third chapter (Orthographic Marker) is illustrated in detail. Red blobs refer to open and blue blobs to closed syllables. (Color figure online)

Phonological awareness refers to the ability to deal with the sound system of a language and to detect, distinguish and manipulate segments of a language like syllables, rimes, or even single sounds [20]. It also includes the adequate perception and processing of prosodic features such as syllable stress. Performance in detecting stress highly correlates with reading and writing skills [7,14,28] and recent research shows that a shortcoming in syllable stress detection in the context of words or sentences is a very strong predictor of dyslexia (e.g., [7,14,21]).

Prosodiya builds on this research and trains stress detection (e.g., Fig. 3a). First, this might boost a child's ability to segment words into relevant components. Second, children learn to focus on relevant areas in words, as orthographic challenges mainly occur in stressed syllables: In the German orthographic system, there is a strong association between stress and vowel length markers – vowel length markers generally occur in stressed syllables [33]. Prosodiya aims at clarifying this association. It helps children to focus on the stressed syllable and to learn how such syllables are spelled. In doing so, it finally leads to a rule-based orthographic spelling training inspired by the empirically evaluated Marbuger Rechtschreibtraining [17].

Little inhabitants called "Kugellichter" ("spherical lights"), kindred to will-o-wisps, guide the children through the world of syllables and orthography and accompany them through the story: The magical land of Prosodiya is haunted by a mysterious and maleficent fog, causing the inhabitant to live in sadness. Only the children can relieve the world from its sorrow. Prosodiya consists of six chapters, each corresponding to a different linguistic or orthographic challenge for which different mini-games were developed. Each chapter is composed of various subchapters, each targeting a specific level of linguistic or orthographic competence. Each subchapter is again composed of various levels. The levels increase in difficulty in that the target words' structures get more complex and the objective(s) of tasks get more challenging. The structure of Prosodiya can be seen in Fig. 2. Each part of the game is introduced and instructed by an interactive tutorial that explains game mechanics, imparts linguistic knowledge, and narrates the story. An integrated learner model aims at adapting the game to their individual proficiency, trying to keep up the learning curve, motivation, and fun. We give a brief introduction into the therapeutic process and its games. A further overview with videos is available at http://youtube.prosodiya.de.

(a) Stress Pattern (b) Open & Closed Syllables (c) Orthographic Marker

(d) Syllable Arcs (e) Embodied Syllable Arcs (f) Spelling

Fig. 3. Exercises of Prosodiya. See http://youtube.prosodiya.com for videos. (Color figure online)

In the first chapter, children develop and improve their awareness for syllable stress. They rebuild stress patterns of audio- and/or visually presented words by dragging and dropping cartoon blobs (Fig. 3a, big, green blobs for stressed and yellow blobs for unstressed syllables). This awareness is extended in the second chapter in which the structure of the stressed syllable is explored (Fig. 3b). They additionally have to decide whether the stressed syllable of a word is open (ends with a long vowel, represented by red blobs) or closed (ends with a consonant that closes the syllable and "squeezes" the vowel, represented by blue blobs).

Processing the structure of the stressed syllables provides a basis for acquiring the complex spelling rules that underlie spelling of long vs. short vowels in the German orthography. In the third chapter, children need to find out the spelling of these structures (Fig. 3c). They acquire knowledge about the rules that underlie the spelling of open and closed syllables by playing various minigames that cover the recognition of each of the special orthographic markers (c.f. [15] for a detailed description). In the fourth chapter, they learn to divide written words into relevant components (e.g., syllables) and thus foster their orthographic representation. The traditional intervention "draw syllable arcs" is commonly used to train syllable analysis. In Prosodiya, this training is enhanced by emphasizing syllable stress (children draw outstanding arcs for stressed syllables). This game is also being developed as a so-called embodied training that uses body movement and gestures. In this version, children speak each syllable clearly and loudly, simultaneously do a sidestep, and swing their writing hand from their left to their right shoulder (Fig. 3e). Fitness trackers with built-in accelerometers are used to record and classify the swung stress pattern. We refer to [16] for a detailed description about this embodied training and its implications.

In the last chapter, children foster their previously acquired knowledge by spelling out words using a predefined set of letters (Fig. 3f). This set of letters can - depending on the difficulty - contain distractors.

Table 1. The questionnaires given to children (C) and parents (P).

	Question	Response options
	Children and Parents	
Q1	C: *How do you like Prosodiya?* P: —"—	awful ○○○○○ great
Q2	C: *Did you enjoy training with Prosodiya?* P: *Did your child enjoy —"—?*	not at all ○○○○○ very much
Q3	C: *Would you like to continue training with Prosodiya?* P: *Would you continue using Prosodiya for your child?*	☐ yes ☐ yes, if new games added ☐ don't know ☐ no
Q4	C: *Do you think Prosodiya helped you to improve your reading and writing skills?* P: *Do you think Prosodiya helped your child to improve their reading and writing skills?*	☐ yes ☐ don't know ☐ no
	Only Children	
Q5	C: *How did you like the graphics?*	awful ○○○○○ great
Q6	C: *How do you like our Kuggelichter?*	not at all ○○○○○ very much
Q7	C: *How did you like the tutorials?*	not at all ○○○○○ very much
Q8	C: *How difficult was Prosodiya?*	very hard ○○○○○ very easy
Q9	C: *Was the task's objective always clear?*	☐ yes ☐ sometimes ☐ no

2.1 Evaluation of a Pilot Study

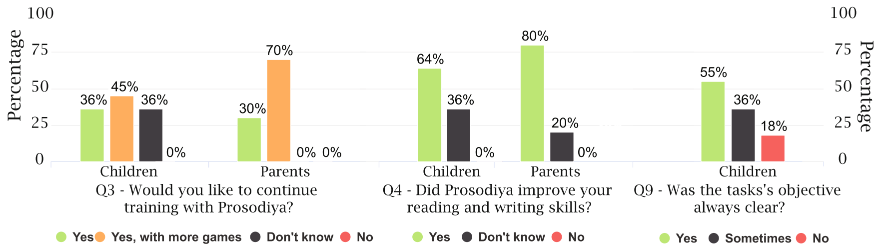
We conducted a pilot study with a preliminary version of Prosodiya in winter of 2015 with 11 dyslexic children from 2nd to 7th grade ($M = 4$, $SD = 1$) aged 7–13 ($M = 9.57$, $SD = 1.63$). Nine of the children were boys and two were girls. The study version contained the first two chapters and a restricted version of the third chapter (different orthographic markers were introduced in one and the same game – not in separate games as in the current version). In total, it consisted of 29 levels and covered 220 words. The children spent an average time of 192.6 min ($SD = 70.69$) training in-game during a period of six weeks.

We used quantitative spelling (DRT 2–5, i.a [26]) and reading tests (SLRT II [25], SLS [23]) to evaluate changes in literacy competence. Two questionnaires (see Table 1) were given to the families after the post-test to assess effects on motivation, enjoyment, and self-efficacy. One was answered by the children and one by their parents. The parents of one child didn't answer the questionnaire.

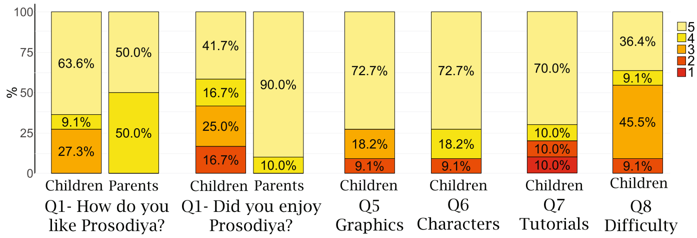
Results

Results of the questionnaires are listed in Figs. 4a and 4b.

Children and parents rated Prosodiya altogether (Q1) with an average of 4.36 and 4.5 out of 5 points. Children enjoyed playing Prosodiya (Q2, $M = 4.0$)



(a) Answers given to the multiple choice questions Q3, Q4, and Q9.



(b) Answers given to the Likert scale questions Q1, Q2, Q5, Q6, Q7, and Q8.

Fig. 4. Results of the questionnaires.

which is in line with the impression their parents reported ($M = 4.5$). Children rated the graphical appearance of the game with an average of 4.36 points (Q5) and its main characters with an average of 4.55 points (Q6). The majority of both children (72%) and parents (100%) would like to continue using Prosodiya (Q3), especially if more games are added (45%, 70%, respectively). Also the majority of both children (64%) and parents (80%) reported that they perceive self-efficacy considering reading and writing skills (Q4).

Children rated the overall difficulty of Prosodiya (Q8) with an average of 3.73 points ($SD = 1.1$) of a scale ranging from 1 (very hard) to 5 (very easy).

The majority of the children (55%) reported that the task's objective was always clear (Q9), 36% reported that it sometimes took some time to figure out the objective and 18% reported that they often had to guess what the task's objective was. One child answered this questions with both of the two latter options resulting in a total response above 100%. The tutorials (Q7) were rated with an average score of 3.82 points ($SD = 0.73$). Two children rated them very low (1 and 2 points) whereas the remaining rated them on average with 4.87 points.

In spelling tests, six children improved their performance in post-tests (DRT 2–5, i.a [26]), only two did not change in performance and one child performed worse. In speeded single word reading post-tests (SLRT II [25]), three children improved, four children did not change in performance and two children performed worse. In speeded reading comprehension tests (SLS [23]), the best improvements were obtained. Almost all children improved with up to 13 points on the fluency scale (a competence level encompasses 9 points). Only one child's performance did not change.

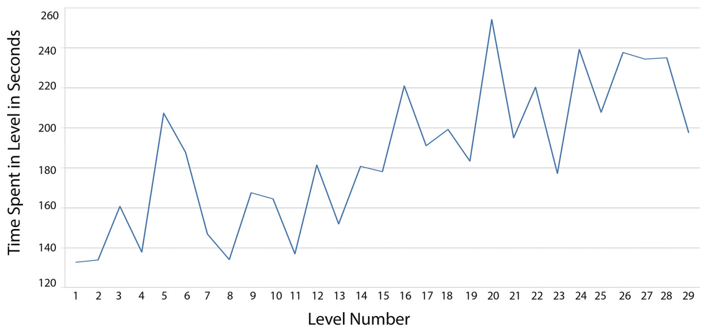


Fig. 5. Average time spent in the different levels. The curve resembles the flow channel used in game design.

Discussion

The results of the questionnaires suggest that both children and parents like Prosodiya (Q1), enjoy spending time playing it (Q2), and are waiting for more content (Q3). The graphical elements of Prosodiya seem to be appealing and match the taste of primary-school children (Q5), which is crucial for serious games [3]. Also the main characters seem to be well received (Q6).

According to the results of Q8, the perceived level of difficulty seems to be adequately between boredom (too easy) and frustration (too hard). Our goal was to keep the children within the flow-channel, a narrow band between boredom and anxiety [9]. In-game measurements of average time spent in a specific level (Fig. 5) indicate the success of our approach. As the number of tasks for each level was constant, the difference of time spent can be attributed to response times for single tasks - which in turn can be used to estimate the proficiency level of a learner [31]. Thus, the in-game metrics indicate that sequences of tension (more time spent, caused by increased difficulty) and relaxation (less time spent, caused by a higher increase of learner proficiency compared to difficulty adjustment) keep the children in a state of flow. This in turn can have positive impact on learning and player's attitude [19].

The reports to Q4 are important to the area of self-efficacy and self-esteem. Boosting both is a central aim of therapeutic intervention [2]. The perceived high self-efficacy reported by children and parents is related to self-awareness of skill increase and actual skill increase [8]. As pointed out by [3], the absence of or little self-awareness of skill increase must be avoided.

We have analysed the reports regarding the questions whether the games were self-explanatory to a satisfactory degree (Q8) and how the tutorials were rated (Q7). Although the majority answered both questions quite positively, the reports also indicate room for improvement. The reasons for negative reports can be that the tutorials didn't communicate the objective of a game well enough or the fact that sometimes the difficulty affecting game mechanics or objectives was slightly increased between levels without informing the children. We therefore derived and applied three changes: First, we refined the tutorials by splitting complex domains and addressing the respective domain in much more detail and including more examples. Second, whenever something changes that could have an impact on the children's answers and behaviour, a brief information is displayed and informs the children of the change and its consequences. Third, we developed in-game tool-tips for each game that explain the game's objective(s).

Overall, in spelling as well as in reading, approximately half of the children could improve their performance, some of them significantly. About a third did not change in performance and two children performed worse. Overall the best improvements were obtained in speeded reading comprehension tests. These results are promising, considering that our pilot's duration was only six weeks. A meta-analysis [13] has shown that interventions with a maximum duration of 12 weeks have only small effect sizes and interventions that last more than 12 weeks have a higher mean effect size. The main objective of the present pilot was primarily to investigate user experience and playability, therefore, duration was set to six weeks. But for future work, focus will be set on therapeutical efficacy and intervention phases will be prolonged.

All in all, the evaluation of the pilot study suggests that the version of Prosodiya at that time reaches a satisfactory level in terms of game and therapeutic design. However, it also highlights the limitations of that version, especially in terms of effects of the game on reading and writing performance. The evaluation of the questionnaires provide evidence that our proposed ICCGD is a valid approach for designing serious games for children. It also highlights the drawbacks of the study version that we addressed in the current version.

Limitations

The pilot study had two major limitations with respect to training effects. First, the children spent significantly less time training with Prosodiya than is recommended by [13]. To find empirical evidence for the effects on reading and writing improvement, the intervention period must be extended and intensified. Second, the games were very limited. We didn't include spelling practice. However, training phonological awareness is assumed only to have an effect on writing performance if combined with spelling exercises (cf. [30]). The crucial core component of spelling was added to the current version of Prosodoiya.

3 Conclusion and Future Work

In summary, we presented a mobile serious game for dyslexic children and the results of a pilot study of a preliminary version of the game. Prosodiya introduces various novelties in this area of research. The main contribution of Prosodiya is its unique focus on syllable stress that we derived reasonably. This focus led to the development of innovative exercises based on empirical findings. The pilot study was conducted with 11 dyslexic children. A central feedback in questionnaires was that the children liked playing the game and that they wanted more levels to keep on playing. Overall, they felt that they could improve their literacy skills. This self-assessment was not represented in quantitative reading and spelling tests in which almost half of the children improved after a period of six weeks, the other half did not show any change in competence, and two children performed worse.

To evaluate the effects of this therapy approach in an improved and most recent version of Prosodiya, a randomized control trial with a waiting control group design is planned starting January 2018. Both a group of dyslexic and unimpaired children from 2nd to 4th grade will practice in total eight weeks with Prosodiya. In the intervention phase, children should practice 20 min per day, 5 days per week. This will result in approximately 800 min of training. Besides the effect of the therapeutic approach on reading and writing, newly introduced elements of gamification will also be investigated.

References

1. Beddington, J., Cooper, C.L., Goswami, U., Huppert, F.A., Jenkins, R., Jones, H.S., Tom, B.L., Sahakian, B.J., Thomas, M., Field, J., Goswami, U., Huppert, F.A., Jenkins, R., Jones, H.S., Kirkwood, T.B.L., Sahakian, B.J., Thomas, S.M.: The mental wealth of nations. Nature **455**(7216), 1057–1060 (2008)
2. Bender, F., Brandelik, K., Jeske, K., Lipka, M., Löffler, C., Mannhaupt, G., Naumann, C.L., Nolte, M., Ricken, G., Rosin, H., Scheerer-Neumann, G., von Aster, M., von Orloff, M.: Die integrative Lerntherapie. Lernen und Lernstörungen **6**(2), 65–73 (2017)
3. Berkling, K., Faller, H., Piertzik, M.: Avoiding failure in modern game design with academic content - a recipe, an anti-pattern and applications thereof. In: Proceedings of the 9th International Conference on Computer Supported Education, pp. 25–36 (2017)
4. Berkling, K., Pflaumer, N., Lavalley, R.: German phonics game using speech synthesis - a longitudinal study about the effect on orthography skills. In: Workshop on Spoken Language Technology for Education (SLaTE), pp. 167–172 (2015)
5. Boyle, E.A., Hainey, T., Connolly, T.M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., Pereira, J.: An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. Comput. Educ. **94**, 178–192 (2016)
6. Bradley, L., Bryant, P.E.: Categorizing sounds and learning to read - a causal connection. Nature **301**(3), 419–421 (1983)
7. Brandelik, K.: Sprachrhythmische Fähigkeiten im Schriftspracherwerb. Südwestdeutscher Verlag für Hochschulschriften (2014)
8. Cleary, T.J., Velardi, B., Schnaidman, B.: Effects of the Self-Regulation Empowerment Program (SREP) on middle school students' strategic skills, self-efficacy, and mathematics achievement. J. Sch. Psychol. **64**, 28–42 (2017)
9. Csikszentmihalyi, M.: Beyond Boredom and Anxiety. The Jossey-Bass Publishers, San Francisco (1975)
10. Daniel, S.S., Walsh, A.K., Goldston, D.B., Arnold, E.M., Reboussin, B.A., Wood, F.B.: Suicidality, school dropout, and reading problems among adolescents. J. Learn. Disabil. **39**(6), 507–514 (2006)
11. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining "Gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011, pp. 9–15. ACM, New York (2011)
12. Fullerton, T.: Game Design Workshop: A Playcentric Approach to Creating Innovative Games, 2nd edn. Morgan Kaufmann, Heidelberg (2008)

13. Galuschka, K., Ise, E., Krick, K., Schulte-Körne, G.: Effectiveness of treatment approaches for children and adolescents with reading disabilities: a meta-analysis of randomized controlled trials. PLoS ONE **9**(2), 1–12 (2014)
14. Goswami, U., Mead, N., Fosker, T., Huss, M., Barnes, L., Leong, V.: Impaired perception of syllable stress in children with dyslexia: a longitudinal study. J. Mem. Lang. **69**(1), 1–17 (2013)
15. Holz, H., Beuttler, B., Brandelik, K., Brandelik, J.: Prosodiya - ein Lernspiel zur Förderung des Sprachrhythmus bei Kindern mit LRS. In: Burghardt, M., Wimmer, R., Wolff, C., Womser-Hacker, C. (eds.) Mensch und Computer 2017 - Tagungsband, pp. 395–398. Gesellschaft für Informatik e.V, Regensburg (2017)
16. Holz, H., Beuttler, B., Kirsch, A.: Bewegungserkennung mit Wearables für Embodied Trainings in Serious Games. In: Burghardt, M., Wimmer, R., Wolff, C., Womser-Hacker, C. (eds.) Mensch und Computer 2017 - Tagungsband, pp. 259–262. Gesellschaft für Informatik e.V, Regensburg (2017)
17. Ise, E., Schulte-Körne, G.: Spelling deficits in dyslexia: evaluation of an orthographic spelling training. Ann. Dyslexia **60**(1), 18–39 (2010)
18. Kast, M., Baschera, G.M., Gross, M., Jäncke, L., Meyer, M.: Computer-based learning of spelling skills in children with and without dyslexia. Ann. Dyslexia **61**(2), 177–200 (2011)
19. Kiili, K., de Freitas, S., Arnab, S., Lainema, T.: The design principles for flow experience in educational games. Proc. Comput. Sci. **15**, 78–91 (2012)
20. Klicpera, C., Schabmann, A., Gasteiger-Klicpera, B.: Legasthenie - LRS, 4th edn. Reinhardt, München (2013)
21. Landerl, K.: Categorization of vowel length in German poor spellers: an orthographically relevant phonological distinction. Appl. Psycholinguist. **24**(4), 523–538 (2003)
22. Liebal, J., Exner, M.: Usability für Kids. Vieweg+Teubner, Wiesbaden (2011)
23. Mayringer, H., Wimmer, H.: SLS 1–4 - Salzburger Lesescreening für die Klassenstufen 1–4. Verlag Hans Huber, Bern (2003)
24. Moll, K., Landerl, K.: Double dissociation between reading and spelling deficits. Sci. Stud. Read. **13**(5), 359–382 (2009)
25. Moll, K., Landerl, K.: SLRT II - Lese- und Rechtschreibtest. Verlag Hans Huber, Bern (2010)
26. Müller, R.: Diagnostischer Rechtschreibtest für 2. Klassen (DRT 2). Beltz, Göttingen (2003)
27. Reuter-Liehr, C.: Behandlung der Lese-Rechtschreibschwäche nach der Grundschulzeit: Anwendung und Überprüfung eines Konzeptes. Zeitschrift für Kinder- und Jugendpsychiatrie **21**(3), 135–147 (1993)
28. Sauter, K., Heller, J., Landerl, K.: Sprachrhythmus und Schriftspracherwerb. Lernen und Lernstörungen **1**(4), 225–239 (2012)
29. Schulte-Körne, G., Remschmidt, H.: Legasthenie - Symptomatik, Diagnostik, Ursachen, Verlauf und Behandlung. Dtsch Arztebl International **100**(7), 396–406 (2003)
30. Schulte-Körne, G., Galuschka, K.: Diagnostik und Behandlung von Kindern und Jugendlichen mit Lese- und / oder Rechtschreibstörung (2015)
31. Sense, F., Behrens, F., Meijer, R.R., van Rijn, H.: An individual's rate of forgetting is stable over time but differs across materials. Top. Cogn. Sci. **8**(1), 305–321 (2016)
32. Snowling, M.J.: Phonological processing and developmental dyslexia. J. Res. Reading **18**(2), 132–138 (1995)
33. Staffeldt, S.: Einführung in die Phonetik, Phonologie und Graphematik des Deutschen Ein Leitfaden für den akademischen Unterricht. Stauffenburg (2010)

A Spinal Column Exergame for Occupational Health Purposes

Sergio Valdivia[1], Robin Blanco[1], Alvaro Uribe[3(✉)], Lina Penuela[1], David Rojas[2], and Bill Kapralos[3]

[1] Universidad Militar Nueva Granada, Bogota, Colombia
sergiodragoon@gmail.com, {u3900208,lina.penuela}@unimilitar.edu.co
[2] The Wilson Centre, Toronto, ON, Canada
davidrojasgualdron@gmail.com
[3] University of Ontario Institute of Technology, Oshawa, ON, Canada
alvaro.j.uribe@ieee.org, bill.kapralos@uoit.ca

Abstract. Sedentary, bad posture, and repetitive movements are the leading cause of several health problems, including spinal column pain, which, can lead to work absenteeism, deterioration of the quality of life, and surgery in extreme cases. Physical activity and exercise can reduce the risk of spinal column problems, and pain. However, sustaining healthy habits, such as exercising, requires motivation and engagement. Exergames are becoming more relevant thanks in part to the use of wearable technologies that provide compelling, engaging, and motivating experiences that can help improve health care. In this work, we present the development of a spinal column occupational health exergame and a study on engagement with two motion tracking technologies, a Microsoft Kinect V2 sensor, and an inertial measurement device. Results indicate, that the inertial measurement unit performs better than Microsoft's Kinect V2, but the game was perceived as more engaging using Microsoft's Kinect V2.

Keywords: Exergame · Posture tracking · 3DUI

1 Introduction

Approximately 83% of the world's population lives with some form of a musculoskeletal disorder (MSD) due to sedentary, bad posture, and repetitive movements as a result of workplace activities, amongst others [1]. Additionally, the ubiquitous use of computing, and mobile devices both in the workplace and in our personal lives, is leading to greater health risks. MSDs may result in work absenteeism, reiterative visits to medical facilities, physiotherapy, and surgery in chronic cases [2]. Currently, physical activity remains the best option to prevent work-related MSDs [3], through typically short and slow-paced upper, lower and back flexion and extension movements, often performed without supervision [1]. Improper execution of physical activity may lead to health problems that could

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 83–92, 2017.
https://doi.org/10.1007/978-3-319-71940-5_8

worsen a MSD. This is particularly the case when the physical activity is performed without supervision, where printed and multimedia guides provide the instructions for performing the exercises, leaving the patient without feedback as to how the exercise is being performed [4].

The monitoring and assessment of physical activity plays an important role in assessing progression and providing adjustments to exercise routines. Traditionally, progress assessment has relied on observation, self-reports, and questionnaires that may not provide an accurate overview of the physical activity progression [5]. In sports, physical activity measurements employ high-end motion capture systems based on infrared cameras and gait laboratories, amongst others [6]. In contrast, wearable technologies and motion tracking systems supported by consumer electronic devices, have arisen as consumer-level solutions that can impact a wider range of the population [7].

Despite the inclusion of newer technologies that can provide physical activity feedback, there still exists an underlying challenge associated with a lack of motivation and engagement [8]. As an example, consider the Nintendo Wii remote control (commonly referred as the Wiimote), a video game controller that provided motion-based interactions by employing an accelerometer and a gyroscope [9]. The Wiimote has been applied in health care settings including, the rehabilitation of patients who suffered a stroke [10], and Parkinson patients [11]. Following the success of the Nintendo Wii, Microsoft launched the first version of the Microsoft Kinect, an image-based sensor that allows body motion tracking through stereo cameras to provide natural interactions within gaming [12]. Similarly to the Wiimote, the Microsoft Kinect has also been employed in health care settings for physiotherapy [13]. Currently, open electronics, 3D printing, and inertial measurement units (IMUs) are providing low-cost, consumer-level wearable motion capture alternatives [14].

The coupling of video games and exercise, whereby playing a video game becomes a form of physical activity, is known as exergames [15]. Exergames may increase engagement [16], reduce pain perception [17], and enhance retention and understanding of the physical activity [18]. Although the application of exergames has been primarily oriented to physiotherapy and rehabilitation, there are additional applications in fitness [19], obesity prevention [20], and treatment and prevention of eye [21], hands [22], and lower/upper limbs occupational health problems [23].

This paper describes the design and development of a spinal column exergame as a tool to promote exercise amongst office workers who spend more than seven hours each day seated in front of a computer. Engagement is provided through carefully chosen game mechanics coupled with the physical activity, while motion interactions are captured with two sensors, a Microsoft Kinect V2 and an IMU to analyze the performance with the exergame and without it.

The remainder of this paper is organized as follows: Sect. 2 describes the game analysis and design from a learning-game mechanics perspective. A description of the experimental design is provided in Sect. 3. Experimental results of the preliminary study conducted to examine the engagement capabilities of the exergame, and a comparison of the movements performed between the exergame and a game-less version using different motion tracking devices are provided in Sect. 4. Concluding remarks and plans for future work are provided in Sect. 5.

2 Development

The development of the exergame requires the analysis of the spinal column to determine the ranges of motion and exercising characteristics that define the design process. The spinal column is comprised of five sections: the cervical, thoracic, lumbar, coccidea, and sacrum [24]. Most common back problems are related to the lumbar area due to damage caused by bad posture, overuse, or exceeding forces [25]. Physical activity focused on flexion and extension frontal rotations can alleviate pain, realign the vertebrate, and strengthen the abdominal muscles [26].

2.1 Game Mechanics and Learning Mechanics

The exergame design process is similar to that of "traditional" (entertainment) video games and follows the same structure where game elements are chosen to achieve the intended goal [27]. However, exergames require additional considerations and care to provide suitable interactions that do not affect the game progress and feedback, and are within the user's mobility capacities. While exergames focus on physical activity, they also present learning opportunities to create awareness while playing. To design our exergame, we followed the exergame design of Sinclair, Hingstone, and Masek where attractiveness, effectiveness, input devices, flow, and event tracking are considered as key parameters [28]. Additionally, to further extend our exergame functionality, we employed the analysis of learning and game mechanics conducted by Arnab et al. from where we chose motivation, ownership, competition, and clear instructions with feedback (see Table 1) [29].

Table 1. Identified game and learning mechanics and implementation description.

Game mechanics	Learning mechanics	Implementation and use
Exercise demonstration	Instructions Demonstration Shadowing	The game instructs the proper execution of a movement
Indication of execution	Feedback Imitation	Proper exercise execution rewards the player
Scores	Motivation Competition	Score is awarded as a result of well performed movements
Goals	Motivation Competition Planning	Game goals provide intrinsic motivation

2.2 Game Design

The game design included the definition of the rules based on the physical activity comprised of paced flexion and extension back movements performed during a minute, game mechanics associated to the effects of movement that allows painting a room accordingly to the amplitude of the movement, and the goal of the exergame, consisting in performing five repetitions and fully paint the room. The exergame was implemented using the Unity3D game engine as it allows cross-platform compilation and compatibility with the Arduino open-source platform (used to acquire motion data from the IMU), and the Microsoft Kinect V2 video-based sensor.

To achieve the motion tracking with the Microsoft Kinect V2, we configured its software development kit, and obtained the position data for the 21 body joints tracked by the sensor. We chose the spine-base and the spine-mid joints to record movements associated with the lumbar and thoracic spinal column areas during flexion and extension. The sensor takes samples on 30 Hz with a latency of 60 ms that can be affected by the hardware characteristics of the computer running the sensor.

Motion tracking with the IMU was achieved by employing an Arduino UNO board, and IMU MPU-9250 nine degrees of freedom sensor (i.e., three accelerometer axis, three gyroscope axis, and three magnetometer axis measurements), with a resolution of 16 bits and configurable measuring range, that supports data sampling at 400 kbps. The IMU was strapped to an elastic band worn by the user. To avoid the use of a tether and possible interference of a wired connection, we designed the Arduino motion capture system to work wirelessly with Bluetooth and with a portable battery.

The exergame scenario is comprised of a white room that will be painted through the user's frontal and lateral spinal column flexion and extension movements obtained from the sensors. To achieve the goal of painting the room and account for the differences amongst users, the range of movement can be customized in the calibration stage, allowing to define the motion thresholds. The objective for the customization is to maintain the users within the game flow area with proper tracking to avoid possible overwhelming difficulties [30]. The game indicates which movements are to be performed through visual cues employing an avatar mimicking the user. To further increase user immersion, the painting color changes according to the executed movements. Figure 1 presents the graphical user interface (GUI), and the virtual avatar mimicking the user that provides visual feedback that may help to correct and improve a movement. It is worth noting that motion tracking with the IMU provides spinal column movements only, thus the arms of the avatar do not move as can be seen from Fig. 1.

3 Experimental Design

We designed and conducted a preliminary experiment to measure the engagement capabilities of the exergame, and to gauge the participant's perceptions

Torso flexion with the IMU

Torso lateral flexion with the IMU

Fig. 1. GUI and virtual avatar motion tracking with the Microsoft Kinect V2 sensor and the IMU.

after playing it, while acquiring motion data employing a Microsoft Kinect V2 sensor and IMU sensor, along with a game-less version. Brockmyer et al. [31] defined engagement loosely as “a generic indicator of game involvement”. The experiment was comprised of the following stages: (i) an explanation phase where the facilitators presented the exergame and the motion capture technologies to the participants, (ii) a calibration stage where the facilitator asked the participants to complete the flexion and extension five times as they would do without any gaming component, (iii) an exercising stage with a duration of one minute where the participants positioned in front of a television and the Microsoft Kinect V2 sensor, started interacting with the GUI through the flexion and extension movements, and received feedback from the game, (iv) participants were presented a IMU-based version of the exergame that required them to perform the exercises with a duration of a minute, and finally, (v) the assessment of the exergame where participants completed the Game Engagement Questionnaire (GEQ) [31].

The GEQ contains 19 questions and provides a psychometrically strong measure of levels of engagement specifically elicited while playing video games [31]. Each question can be responded to with one of three options: (i) No, (ii) Maybe, and (iii) Yes, where each option is assigned a numerical value of -1, 0, and 1 respectively. The questions of the GEQ are grouped into the four categories of (i) immersion, (ii) presence, (iii) flow, and (iv) absorption, where -1 represents the lowest engagement perception, while 1 the maximum one, and any value in-between indicates that further work is required to increase the engagement.

3.1 Participants

We invited ten participants to use the spinal column exergame, of whom five used the version of the game that employed the Microsoft Kinect V2 sensor, and five used the version of the game that employed the IMU. Additionally, each participant performed the physical activity with a game-less version of the exergame, while maintaining the same motion capture device.

All participants confirmed working more than seven hours in front of a computer and expressed not exercising their backs while acknowledging the importance to do so.

4 Results

4.1 Microsoft Kinect V2 Data Capture

Participants who played the exergame with the Microsoft Kinect V2 sensor reported a minimum average flexion of 15.4° and a maximum average flexion of 75.6. Figure 2 presents a sample from one participant flexion and extension execution with and without the exergame. It can be observed from the captured data that without the exergame the participant executed the movements as best as possible, but over-exceeding the flexion tracking range. However, when participants focused on performing the exercise as fast as possible without the exergame, the Microsoft Kinect V2 measurements indicate data loss due to the inability of the sensor to keep track of rapid movements in detail. This was corrected with the exergame, as the movements were presented in a slower-paced form that allowed acquiring more accurate information. Results indicate that the use of the exergame allows the participants to perform the movements more consistently and with similar amplitude or angle range, as can be seen in Fig. 2.

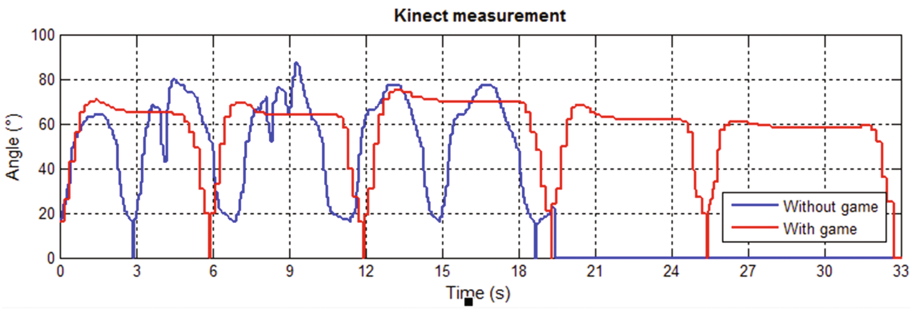


Fig. 2. Microsoft Kinect V2 spinal column motion tracking comparison with and without the exergame. Blue - motion tracking without the game. Red - motion tracking with the game. (Color figure online)

4.2 IMU Data Capture

The participants who used the exergame with the IMU presented a 5.7° minimum flexion and a maximum of 71.2° on average. Motion-tracked data with the complementary filter programmed in Unity to process the Arduino board signals resulted in a minimum noise level under one degree which was disregarded, proving adequate for motion tracking purposes. Similar to the measurements obtained with the Microsoft Kinect V2 sensor, Fig. 3 presents the motion tracked data of

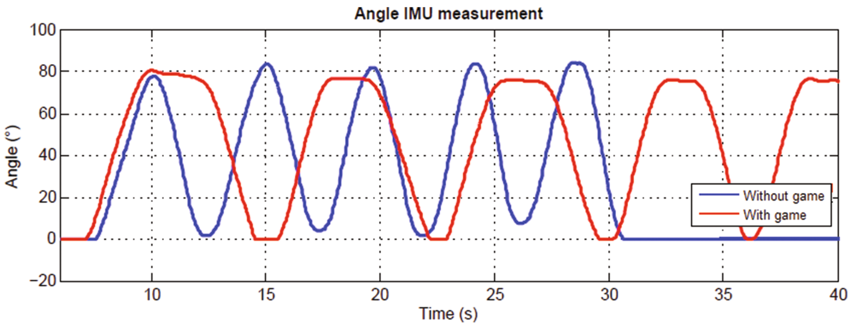


Fig. 3. IMU spinal column motion tracking comparison with and without the exergame. Blue - motion tracking without the game. Red - motion tracking with the game. (Color figure online)

a single user performing the flexion and extension movement with and without the exergame.

Interestingly, the data captured with the IMU allows observing a paced execution with flexion ranges that allow keeping the user within the visibility of the sensor. Additionally, although in this case the participant performed the movements faster without the exergame, the data is more accurate and adequate to assess the motion range, which, in comparison with the Microsoft Kinect V2 sensor provides more reliability.

4.3 GEQ Analysis

The results of the GEQ provided engagement insights while playing the game with the Microsoft Kinect V2 sensor and the IMU. As previously described, the GEQ includes four categories for measuring engagement including immersion, presence, flow, and absorption [31]. Although several definitions can be found for these four terms, here we present the definitions described by Brockmyer et al. [31]: (i) immersion is the experience of becoming engaged in a virtual experience while being aware of one's surroundings, (ii) presence is being in a normal state of consciousness while being inside a mediated virtual environment, (iii) flow describes the feelings of enjoyment as a result of the interaction balance between skill and challenge, and finally (iv) absorption is the total engagement within the virtual scenario that causes an altered state of consciousness.

The GEQ results (as presented in Table 2), indicate that those who played the game where the tracking was accomplished using the Microsoft Kinect V2 sensor perceived it as being more immersive than with the game that employed the IMU, even though these obtained the same GEQ score. We believe the increase in immersion was the result of the Microsoft Kinect V2 sensor motion tracking capabilities, as it can detect body movements from 26 joints resulting in the avatar mirroring the participant's movements more naturally, while the IMU was configured to only capture the back flexion and extension and all body parts remain static. Presence was higher for the exergame that employed the IMU, and we believe this occurred as a result of the lag of movements present while using the Microsoft Kinect V2 sensor. The flow was overall poorly rated with both tracking technologies indicating that the relation between skill and challenge needs to be improved to take better advantage of the motion-based interaction with the game mechanics. Finally, absorption presented the lowest of the four scores, indicating the need of improvements in taking further advantage of the game elements to alter the state of consciousness of the players.

Table 2. GEQ gathered data for all four categories (*IMU Captured data)

Participant	1	2	3	4	5	Mean
Immersion	1	1	1	1	1	1
	1*	0*	−1*	1*	0*	0.2*
Presence	0.25	0	0.5	0	0.5	0.25
	0.75*	0.5*	0.5*	0.5*	1*	0.65*
Flow	0.23	−0.34	0	−0.12	0.12	−0.02
	−0.12*	0.34*	0.45*	0*	0.56*	0.24*
Absorption	−0.4	−0.2	−0.6	−1	−0.4	−0.52
	−0.6*	0.2*	0.6*	0.2*	0.6*	0.2*

5 Conclusions and Discussion

Here we have presented the design and development of an exergame aimed at low-back care. We employed the Microsoft Kinect V2 sensor and a IMU with an Arduino Uno board to obtain measurable spinal column flexion and extension data from a series of five repetitions. To gauge the engagement perception of the exergame, we compared participant performance with the Microsoft Kinect V2 sensor and an IMU. The GEQ provided relevant information that can help us improve the exergame. From the motion captured data, it was interesting to observe the differences from both sensors, and how the participants approached the goal of exercising with and without the gaming component. With respect to the sensors, the Microsoft Kinect V2 sensor provided a more immersive experience and its compatibility with numerous commercial exergames adds expectation to the users. However, the IMU excelled at better capturing the ranges of movement, proving to be a better device for motion tracking. Additionally, the IMU and the Arduino provided an interesting complement that could be integrated to commercial exergames to provide further data that can assist users in better understanding their progress during the physical activity.

Future work will focus on improvements regarding immersion, presence, flow, and absorption to maximize the effects of the game mechanics. In addition, we will conduct research on improving the acquisition of the data towards developing a framework that allows multiple physiological measurements, increase the number of participants, and analyses over time to understand the possible effects of time of using the exergame.

Acknowledgements. The authors would like to thank Universidad Militar Nueva Granada for funding Project ING-2377.

References

1. Luttmann, A., Jager, M., Griefahn, B., Caffier, G., Liebers, F., Organization, W.H., et al.: Preventing musculoskeletal disorders in the workplace (2003)
2. World Health Organization: WHO — Occupational and work-related diseases. WHO (2012)
3. Brownson, R.C., Hoehner, C.M., Day, K., Forsyth, A., Sallis, J.F.: Measuring the built environment for physical activity: state of the science. Am. J. Prev. Med. **36**(4), S99–S123 (2009)
4. Cromie, J.E., Robertson, V.J., Best, M.O.: Occupational health and safety in physiotherapy: Guidelines for practice. Aust. J. Physiotherapy **47**, 43–51 (2001)
5. Craig, C.L., Marshall, A.L., Sjöström, M., Bauman, A.E., Booth, M.L., Ainsworth, B.E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J.F., et al.: International physical activity questionnaire: 12-country reliability and validity. Med. Sci. Sports Exerc. **35**(8), 1381–1395 (2003)
6. Bassett, D.R.: Validity of four motion sensors in measuring moderate intensity physical activity. Med. Sci. Sports Exerc. **32**(9), S471–S480 (2000)
7. Rosenberger, M.E., Buman, M.P., Haskell, W.L., McConnell, M.V., Carstensen, L.L.: Twenty-four hours of sleep, sedentary behavior, and physical activity with nine wearable devices. Med. Sci. Sports Exerc. **48**(3), 457–465 (2016)
8. Reichert, F.F., Barros, A.J., Domingues, M.R., Hallal, P.C.: The role of perceived personal barriers to engagement in leisure-time physical activity. Am. J. Public Health **97**(3), 515–519 (2007)
9. Wingrave, C.A., Williamson, B., Varcholik, P.D., Rose, J., Miller, A., Charbonneau, E., Bott, J., LaViola Jr., J.J.: The wiimote and beyond: Spatially convenient devices for 3d user interfaces. IEEE Comput. Graphics Appl. **30**(2), 71–85 (2010)
10. Brosnan, S.: The potential of wii-rehabilitation for persons recovering from acute stroke. Phys. Disabil. **32**(1), 1–3 (2009)
11. Herz, N.B., Mehta, S.H., Sethi, K.D., Jackson, P., Hall, P., Morgan, J.C.: Nintendo wii rehabilitation (Wii-hab) provides benefits in parkinson's disease. Parkinsonism Relat. Disord. **19**(11), 1039–1042 (2013)
12. Zhang, H.: Head-mounted display-based intuitive virtual reality training system for the mining industry. Int. J. Min. Sci. Technol. **27**(4), 717–722 (2017). Special Issue on Advances in Mine Safety Science and Engineering
13. Luna-Oliva, L., Ortiz-Gutiérrez, R.M., Cano-de la Cuerda, R., Piédrola, R.M., Alguacil-Diego, I.M., Sánchez-Camarero, C., C, M.C.: Kinect Xbox 360 as a therapeutic modality for children with cerebral palsy in a school environment: a preliminary study. NeuroRehabilitation **33**(4), 513–521 (2013)
14. Case, M.A., Burwick, H.A., Volpp, K.G., Patel, M.S.: Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA **313**(6), 625–626 (2015)
15. Pirovano, M., Surer, E., Mainetti, R., Lanzi, P.L., Borghese, N.A.: Exergaming and rehabilitation: A methodology for the design of effective and safe therapeutic exergames. Entertainment Comput. **14**, 55–65 (2016)

16. Campbell, R., Evans, M., Tucker, M., Quilty, B., Dieppe, P., Donovan, J.: Why don't patients do their exercises? Understanding non-compliance with physiotherapy in patients with osteoarthritis of the knee. J. Epidemiol. Community Health **55**(2), 132–138 (2001)
17. Malloy, K.M., Milling, L.S.: The effectiveness of virtual reality distraction for pain reduction: a systematic review. Clin. Psychol. Rev. **30**(8), 1011–1018 (2010)
18. Macvean, A., Robertson, J.: Understanding exergame users' physical activity, motivation and behavior over time. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1251–1260. ACM (2013)
19. Brauner, P., Calero Valdez, A., Schroeder, U., Ziefle, M.: Increase physical fitness and create health awareness through exergames and gamification. In: Holzinger, A., Ziefle, M., Hitz, M., Debevc, M. (eds.) SouthCHI 2013. LNCS, vol. 7946, pp. 349–362. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39062-3_22
20. Vander Schee, C.J., Boyles, D.: exergaming, corporate interests and the crisis discourse of childhood obesity. Sport Educ. Soc. **15**(2), 169–185 (2010)
21. Navia, M., Uribe-Quevedo, A.: Development of an eye tracking occupational health game. In: 1st International Workshop on Assistive Technologies (2015)
22. Ortiz, S., Uribe-Quevedo, A., Kapralos, B.: Hand VR exergame for occupational health care. In: MMVR, pp. 281–284 (2016)
23. Ramos-Montilla, E., Uribe-Quevedo, A.: Development of an open electronics user inerface for lower member occupational health care exergaming. In: Stephanidis, C. (ed.) HCI 2015. CCIS, vol. 529, pp. 478–483. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-21383-5_80
24. Gianino, J.M., Paice, J.A., York, M.M.: Spinal cord anatomy. Intrathecal Drug Therapy for Spasticity and Pain, pp. 3–14. Springer, New York (1996)
25. Adams, M.A.M.A.: The biomechanics of back pain. Churchill Livingstone/Elsevier (2013)
26. Adams, M.A., Bogduk, N., Burton, K., Dolan, P.: The Biomechanics of Back Pain. Elsevier Health Sciences (2012)
27. Fullerton, T.: Game Design Workshop: A Playcentric Approach to Creating Innovative Games. CRC Press (2014)
28. Sinclair, J., Hingston, P., Masek, M.: Considerations for the design of exergames. In: Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia, pp. 289–295. ACM (2007)
29. Arnab, S., Lim, T., Carvalho, M.B., Bellotti, F., Freitas, S., Louchart, S., Suttie, N., Berta, R., De Gloria, A.: Mapping learning and game mechanics for serious games analysis. Br. J. Educ. Technol. **46**(2), 391–411 (2015)
30. Cowley, B., Charles, D., Black, M., Hickey, R.: Toward an understanding of flow in video games. Comput. Entertainment (CIE) **6**(2) (2008). 20
31. Brockmyer, J.H., Fox, C.M., Curtiss, K.A., McBroom, E., Burkhart, K.M., Pidruzny, J.N.: The development of the game engagement questionnaire: A measure of engagement in video game-playing. J. Exp. Soc. Psychol. **45**(4), 624–634 (2009)

Gamification

Investigating Motivation in Gamification

Results from an Experimental Pilot Study

Peter Bußwolder[1(✉)] and Andreas Gebhardt[2]

[1] Chair of Management Accounting, School of Business and Economics, RWTH Aachen University, Templergraben 64, Aachen, Germany
busswolder@controlling.rwth-aachen.de
[2] Peterstraße 82, Baesweiler, Germany
gebhardt.andreas@rwth-aachen.de

Abstract. Gamification has drawn increasing attention in the last years. The goal is to motivate people and to increase engagement for tasks. However, the empirical research remains limited, due to the unclear definition of the term gamification and an underdeveloped theory. The research at hand tries to narrow this gap by conducting a laboratory experiment to test the effects of gamification on the performance of subjects in a rather boring task. Beyond this, we tried to capture the motivation of the subjects during and after the experiment using different questionnaires. This was done to provide deeper insights into the interim step between gamification and performance. We could identify some significant motivational effects, though the performance was not increased. The results indicate that different gamification elements do not cumulate their influence. Experiments and the measurement of motivational effects of gamification elements provide an interesting research opportunity for future research.

Keywords: Gamification · Experiment · Motivation · Flow

1 Introduction

This paper is concerned with the influence of game-like elements on activities that usually do not show any inherent incentive on itself. This approach is known in literature as gamification and has gained increasing attention from practitioners and researchers alike in the last years [1]. Since then, this approach has been predicted a great potential for implementation in many areas of everyday life. This potential is reflected in the Gartner-Hype-Cycle of 2011, which received the term as a future trend of economics [2]. The concept's practical implementation has found its way into many areas, for example in the fields of customer loyalty, health promotion, social networks, crowdsourced science or computer-assisted learning environments in schools [3, 4, 5].

However, previous recommendations for implementing the gamification approach give heterogeneous advice which and how many game-like elements should be used to create a motivational stimulating design. Frequently, literature only provides a list of possible elements. While Kapp [6] assumes that more than three elements must be present, Werbach and Hunter [7] propose the use of points, awards and a high score as sufficient. This is also seen critically and termed as "pointification" [4, 8]. Reeves and

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 95–104, 2017.
https://doi.org/10.1007/978-3-319-71940-5_9

Read [9], however, suggest the use of ten different elements. Hence, this paper investigates whether the use of the gamification approach has the general potential of raising motivation to perform a task, and whether the various game-like elements do have different motivational effects. Through answering these key questions this paper is trying to shed light on the cumulative use of game-like elements and their influence on motivation compared to single elements. A laboratory experiment was conducted to test the stated research questions. The results provide insights into the effect of gamification on performance and especially on motivation and engagement. This focus on motivation allows opening the black box of gamification and analyzing motivation as the moderator between gamification and increased performance.

The rest of the paper is organized as follows. In the next section, the term gamification is discussed and the relevant literature concerning motivation is reviewed. Section 3 provides the experimental design. In Sect. 4 the results are analyzed. Section 5 finally summarizes and discusses the findings.

2 Theory

Due to the topicality of the approach, many different concepts try to cover the term gamification [10]. One of the first definitions was coined by Detering et al. [11, p. 9]: "Gamification is the use of game design elements in non-game contexts". Kapp [6, p. 9] worked out a more comprehensive explanation of the term, defining it as follows: "Gamification is using game-based mechanics, aesthetics, and game thinking to engage people motivate action, promote learning and solve problems". It is the essential core of both definitions' explaining that the thought of the game is transferred to an everyday context which has nothing to do with gaming, causing a change in behavior. With the gamification approach, this goal should be achieved by the use of single game elements [11] or "Gamification is using game-based mechanics […]" [6, p. 9].

To some extent, the game-like elements form the framework of the gamification approach. A concluding list of all potential elements doesn't seem possible due to the amount [11]. Roughly, it can be differentiated between "self" and "social" elements. The "self" elements include, for example, points, awards and levels, time and time restrictions. These elements take on the role of reshaping an activity in such a way that the player engages in a competition with himself. The "social" elements, on the other hand, for example a ranking or a background story, act in a public space. By using a ranking, a player engages in a competition with his environment [12]. Furthermore, the aesthetics or rather the design can be used as another game-like element. This neither belongs to the "self" nor the "social" elements, but forms a category of its own, as included in the definition by Kapp [6] as a separate explaining element of the gamification approach. Ultimately, the approach has the objective of exerting a positive motivational influence over a person's behavior. By this means, even an unattractive task shall gain appeal and shall be executed with commitment [13].

Precisely this motivational effect is to be examined in this paper. Often, tasks are carried out by individuals only because they offer a substantial motivational appeal, even with the impending threat of negative consequences [14]. Taking up this phenomenon, motivational psychology examines the individual motives behind performed actions. According to Rheinberg [15], two sources of incentives – "action" and "purpose" incentives – are responsible for performing an action. It is the characteristic of an action incentive that the motivation for performing the action lies in the action itself. One performs the action for its own sake. This form of motivation is different from purpose centered forms [15]. Here, a cognitive process of evaluation of the results of an action takes place, with the individual making strategic considerations about his activities. This way, an activity can be performed even if the task itself generates no joy or fun [14]. Within this paper, action incentives will be focused. This limitation is reasonable, since a task which already causes joy during its performance doesn't need further measures to be carried out. Consequently, a positive focus on action offers a solid foundation for further forms of motivation [15].

An especially action focused motivation is the flow experience by Csikszentmihalyi [14, 16]. Early on, Csikszentmihalyi [17] recognized a connection to games in general. Here a connection to the gamification approach emerges. A flow experience means "[…] being absorbed, without reflection, in a smooth-running task, which is under control despite high strain" [18, p. 161]. Thus, the characteristic of a flow experience is the merging of action and consciousness of a person. In a way, it is acted like in a state of trance. The experience of this state is only possible if the task's challenge level is within the actor's scope of performance [16]. Primarily, a person's individual feeling is responsible for the fit with the scope of performance. The person has to individually assess whether they feel in control of the action, regardless of the actual difficulty of the task [15, 16]. Therefore, a flow experience will typically be experienced only in a flow channel, forming in the field of tension between challenge and skills [19]. With the Flow Short Scale, Rheinberg and Vollmeyer [18] developed a survey instrument capable of considering all elemental components of flow experience while still allowing to measure the flow experience over the course of any task. The task must not be interrupted for too long during the answering of the scale. To meet this obligation, the scale only contains thirteen items in total, measured on a seven-points Likert scale ranging from "does not apply" to "does apply" [20]. The Flow Short Scale was used within the framework of the Experience Sampling Method (ESM) [15, 21, 22]. With this method, test subjects assessed their motivational experience during the task with a standardized Flow Short Scale after a signal was given, while not interrupting the task for long.

To test if the gamification approach does set a motivational incentive, the Incentive Focus Scale by Rheinberg [23, 24] is being used. The scale measures if a person is more focused on the enjoyment of performing an activity or rather on the results value. The situational context (work or recreation) has an influence on this relationship.

Depending on the treatment was determined, additionally, if it was important to the test subjects to reach a score as high as possible, respectively a good placing in the ranking. Moreover, it was possible to gain further data regarding the behavior of test subjects in dependence to the respective treatment, due to the experimental design. This includes, for example, the time the test subjects took to complete the experiment or the performance during the experiment.

3 Method

To test the stated research questions, a laboratory experiment was conducted. A two-factorial between-subjects design was used, resulting in four treatments: the control treatment, a treatment with either points or a high score and a complete treatment with both points and a high score. An interactive computer program was developed to serve as the underlying task. A screenshot of the user interface can be seen in Fig. 1.

Fig. 1. General user interface (German original)

In total, 55 students from a major German university participated in the experiment. Of these students 33 were male and 22 female. The average age was 24 years (SD = 2.96). The field of study was mostly business, engineering or industrial engineering. We decided for a fixed payment of 5 € (around $ 5.30) to avoid additional motivational effects, as we set the focus of the motivational effects on the gamification elements [25]. A comparison of the treatment groups did not show any significant differences in relation to the age of the subjects. Fisher's exact test was used to test the independence of the categorical data of gender and field of study [26]. Both null hypotheses that the variables are stochastically independent from the treatment groups cannot be rejected ($p > .62$, respectively $p > .19$, Fisher's exact test).

All subjects were randomly assigned to one of the four treatments. They were placed in front of a computer and were given verbal instructions, which were also available in written form. Furthermore, a short guidance for the task was displayed at the beginning of the computer program. The task consisted of a simple real effort task, which did not require any previous knowledge. The goal was to count the number of appearances of a symbol in a matrix of similar but distinct symbols within a specific set of time. This was repeated 15 times in which the task was getting more difficult as the differences between the correct and wrong symbols got subtler. It is important to note, that the task is rather boring and does not provide any motivation in itself [27].

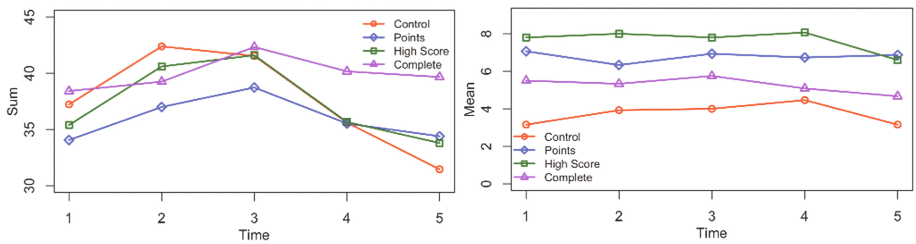
In the control treatment, no further gamification elements besides levels, time pressure and general game aesthetics were used. These were the same in all treatments to provide a general gamefulness [1, 4, 11]. In the point treatment, subjects were awarded with points in accordance to the correctness of their answers, while in the high score treatment they were ranked relative to their success into a list. Both gamification elements, points and high score, were used in the complete treatment. The whole experiment lasted around 30 min.

After every three rounds, the participants were asked to answer the questions of the Flow Short Scale[1], to measure the motivation during the experiment [20]. These interruptions made use of the experience sampling method (ESM), in which participants respond to a short questionnaire while the minimal involvement reduces possible biases [21, 22].

At the end of the experiment, the participants filled out the Incentive Focus Scale [23] as well as several general questions about their experience in the experiment (e.g. was the game fun; would the subject replay it), their gender, age and field of study. After this, they were paid and left.

4 Results

Flow was measured with the Flow Short Scale by [20]. Each of the ten items was measured on a 7-point Likert scale. Additionally, the perceived importance was measured with three items. The responses for each subscale were summed up. The participants answered the Flow Short Scale five times during the experiment, once after three rounds. Figure 2 presents the results for flow (a) and importance (b) for all five times of measurement split by treatment.



(a) Flow (b) Importance

Fig. 2. Results of the flow short scale for flow (a) and importance (b)

The values for Cronbach's Alpha are $\alpha = .973$ for flow and $\alpha = .976$ for importance, respectively. The correlation is positive with $r = .260$ and highly significant ($p < .001$). This is comparable to the literature [20, 28].

[1] All questions and the corresponding statistics can be obtained from the authors.

All four treatments have a similar pattern for flow and importance. Only the Control and Complete treatment show distinct patterns, as the Control treatment has a higher variation and the Complete treatment a lower one. The treatments present only a minor variation in the perception of the importance throughout the experiment, but take different levels of importance. A simple linear model shows a significant effect of the High Score treatment compared to the Control treatment, while flow does not show a significant effect (Model 1 and 2, Table 1). Both flow and importance were analyzed over all five measurement points.

Table 1. Linear models

Variables	Model 1[a]	Model 2[b]	Model 3[c]	Model 4[d]
Intercept	18.692** (6.976)	188.231*** (16.321)	17.308*** (1.058)	14.539*** (0.975)
Treatment points	15.241 (9.531)	−8.497 (22.299)	−0.174 (1.446)	1.528 (1.332)
Treatment high score	19.574* (9.531)	−1.164 (22.299)	−0.374 (1.446)	2.328° (1.332)
Treatment complete	7.641 (10.069)	11.603 (23.558)	−0.224 (1.527)	0.462 (1.407)
Adjusted R^2	0.033	−0.043	−0.057	0.013
F	1.623	0.262	0.023	1.237
Observations	55	55	55	55
Number of parameters	4	4	4	4

Note: *Unstandardized coefficients with standard errors in parentheses*
° p < .1; ∗ p < 0.05; ∗∗ p < 0.01; ∗∗∗ p < 0.001; [a] Dependent variable is sum of importance over all five measurement points; [b] Dependent variable is sum of flow over all five measurement points; [c] Dependent variable is Activity Centering (AC); [d] Dependent variable is Purpose Centering (PC)

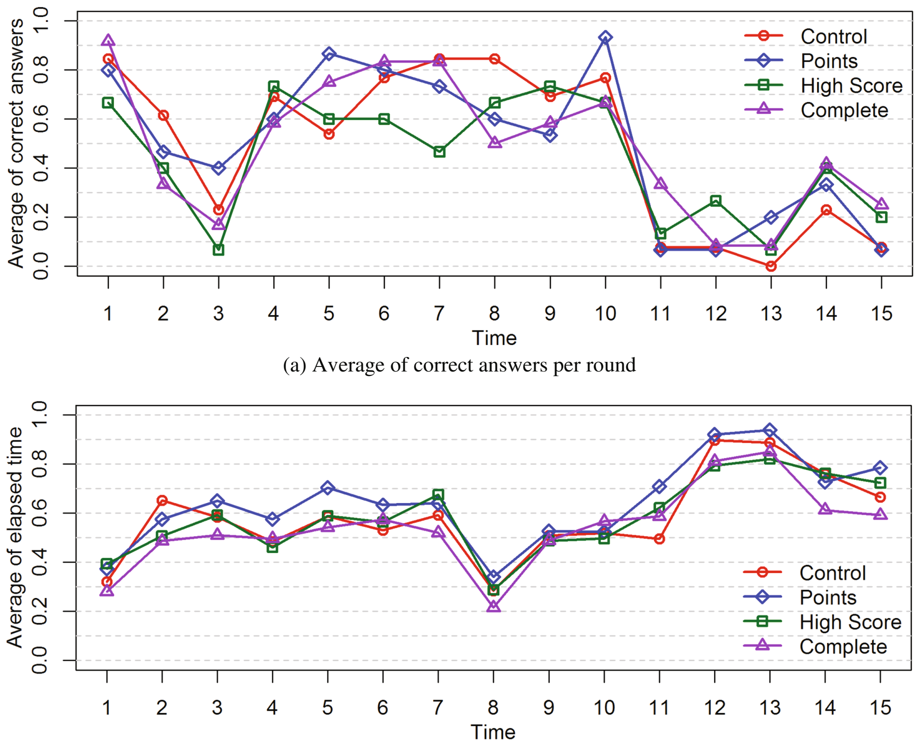
The second scale used in this study is the Incentive Focus Scale [23]. It measures if subjects orient themselves according to the incentive of the task itself or to the accessible outcome (activity vs. purpose centered). Both elements are influenced by the context. For example, in a professional environment the outcome can be of more importance, while during leisure time the activity orientation might be stronger. The scale consists out of 20 questions, 10 for each focus, which are opposed but not completely distinct. People can be activity as well as purpose centered. The subjects evaluated each question on a 4-point Likert scale. The values are summed up to come to a value for activity (AC) and purpose centering (PC). Table 2 provides the results of the Incentive Focus Scale. Using the Shapiro-Wilk-Test to test for normality, PC shows a normal distribution ($W = .959$, $p < .059$) while AC is not normal distributed ($W = .917$, $p > .0001$). This is comparable to previous findings [24]. Not comparable are the low values for internal consistency ($\alpha_{AC} = .515$, $\alpha_{PC} = .572$) and the insignificant correlation ($r = -.140$, $p < .308$).

Table 2. Statistics for Incentive Focus Scale by treatments

	Treatment			
Item	Control	Points	High Score	Complete
Average & standard deviation				
Activity centered	17.308 (3.838)	17.133 (3.357)	16.933 (3.807)	17.083 (4.316)
Purpose centered	14.538 (3.357)	16.067 (3.732)	16.867 (3.796)	15.000 (2.984)

Activity centering does not show any significant treatment effects, while the High Score treatment has a weakly significant effect on purpose centering (Model 3 and 4, Table 1). Here once again does this treatment exhibit a motivational effect. This does not only show up relative within the experiment, but also in comparison to a normative standard. A simulated t-test with the characteristic values stated by Rheinberg et al. [24] for PC, indicates a significant difference to the High Score treatment ($t(14.96) = 2.384$, $p < .033$), indicating that subjects in this treatment were more focused on being successful.

Two outcome measures were considered in this study. These are the distance to the correct answer on the one hand and the elapsed time in percent on the other. The distance to the correct answer is calculated by subtracting the correct answer in a



(a) Average of correct answers per round

(b) Average of elapsed time per round

Fig. 3. Average of correct answers (a) and elapsed time (b) for each round

specific level from the given answer by the subject. A distance of zero indicates a correct answer, while for example a distance of +1 means that the subject deviated by one from the right answer. The time measure is defined by the percentage of the elapsed time in a specific round in relation to the available time. Figure 3 presents the course of the correct answers (a) as well as of the elapsed time (b) for each of the 15 rounds.

The number of correct answers shows a sharp drop after the tenth round afterwards it stays at a low level. It also shows a single drop during the third round. The required time shows an increase over the course of the experiment, except for the eighth round. This indicates an increasing difficulty of the last five rounds. This is supported by the flow experience, which is decreasing for the fourth and fifth measurement.

Statistically, correct answers and elapsed time have a strong negative and highly significant relationship throughout the experiment ($r_{Answer,Time} = -.644$, $p < .001$). This is only logical as the round gets more difficult, more of the provided time is needed. Furthermore, the results for flow for the last three rounds are significantly related to both answers and time ($r_{Flow,Answers} = .248$, $p < .068$; $r_{Flow,Time} = -.298$, $p < .027$).

5 Discussion

The objective of this study was to look into the motivational foundation of gamification. As much of the previous work in this area is conceptual or practically oriented [29], we decided to develop an experiment to empirically analyze the effect of different gamification elements and their combination on performance and motivation.

Our results do not show a significant increase in performance related to gamification. The quality of answers provided by the participants in the different levels do not show significant differences when it comes to points, a high score or both combined. Several explanations are possible. First, the results fall in line with previous research, that partly shows a positive impact of gamification but also a fairly amount neutral or even negative results [4, 30]. This would indicate a weak effect of it in general or the need for further research to develop a deeper understanding of gamification. It is necessary to understand the interaction of game elements, aesthetics and design patterns to develop more sophisticated gamified systems. The elements in this study were selected as they are regularly used in practice due to their simple usage. This is sometimes critically labeled as “Pointification” [4, 8]. Second, as the experiment involves some elements across all treatments, e.g. general game aesthetics or levels, the elements which are altered – points and high score – do not provide further motivational effects on the subjects. Third, as this is mere a pilot study, the number of participants is rather low. This low number may not allow to statistically account for more subtle effects. Further experimental studies are necessary to remedy this limitation.

Again, like the literature, we found motivational effects [4]. Here, especially the flow concept is of interest as it is measured at several points during the experiment using the experience sampling method. Flow and perceived importance provide insights into the differences between points and high score, even though the effects are statistically small. Also, a cumulative effect cannot be supported. The other measures of motivation partly support these findings. This is acceptable, as we involved several measures of motivation to test them for usage in the field of gamification.

Besides providing empirical findings on gamification from a laboratory study, these measures are of particular interest for researchers. Gamification acts through motivating the participants of a system. Deeper insights into these motivational effects are necessary. Also, laboratory experiments represent an interesting research method to test theory. Currently, case studies and practical applications are prevailing [29]. Experiments allow the strict control of treatment factors. They might seem artificial, but in an early state of research, like in gamification, they can provide important insights [31].

Future Research may test more gamification elements and especially more theoretical frameworks. Potential frameworks were proposed for example by Werbach and Hunter [7], Blohm and Leimeister [5] or Aparicio et al. [32]. This can help to develop a sound understanding of gamification to foster practical implications in marketing, health or education. It is the opinion of the authors that empirical work through laboratory experiments to analyze the step between gamification and performance – motivation – is a key to understand gamification and its related concepts. This study involved several motivational measures, forming a start point for further research.

Funding. This research received funding from the fund for quality improvement of the state North Rhine-Westphalia (Germany). The authors declare that they have no conflict of interest.

References

1. Darejeh, A., Salim, S.: Gamification solution to enhance software user engagement - a systematic review. Int. J. Hum.-Comput. Interact. **32**, 613–642 (2016)
2. Muntean, C.: Raising engagement in e-learning through gamification. In: Proceedings of the 6th International Conference on Virtual Learning ICVL, pp. 323–329 (2011)
3. Deterding, S.: Gamification: designing for motivation. Interactions **19**, 14–17 (2012)
4. Seaborn, K., Fels, D.: Gamification in theory and action: A survey. Int. J. Hum. Comput. Stud. **74**, 14–31 (2015)
5. Blohm, I., Leimeister, J.: Gamification: Gestaltung IT-basierter Zusatzdienstleistungen zur Motivationsunterstützung und Verhaltensänderung. Wirtschaftsinformatik **55**, 275–278 (2013)
6. Kapp, K.: The Gamification of Learning and Instruction: Games-Based Methods and Strategies for Training and Education. Wiley, San Francisco (2012)
7. Werbach, K., Hunter, D.: For the Win: How Game Thinking Can Revolutionize Your Business. Wharton Digital Press, New York (2012)
8. Richter, G., Raban, D., Rafaeli, S.: Studying gamification: the effect of rewards and incentives on motivation. In: Reiners, T., Wood, L.C. (eds.) Gamification in Education and Business, pp. 21–46. Springer, Cham (2015)
9. Reeves, B., Read, J.L.: Total engagement: Using games and virtual worlds to change the way people work and businesses compete. Harvard Business School Press, Boston (2009)
10. Miller, C.: The gamification of education. Dev. Bus. Simul. Experiantial Learn. **40**, 196–200 (2013)
11. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining "Gamification". In: MindTrek 2011 Proceedings of the 15th International Academic MindTrek Conference, pp. 9–15 (2011)

12. Huang, W.H.Y., Soman, D.: A Practitioner's Guide to Gamification of Education. Rotman School of Management, Toronto (2013)
13. Stampfl, N.: Die verspielte Gesellschaft: Gamification oder Leben im Zeitalter des Computerspiels. Heise-Verlag, Hannover (2012)
14. Engeser, S., Vollmeyer, R.: Tätigkeitsanreize und Flow-Erleben. In: Vollmeyer, R., Brunstein, J.C. (eds.) Motivationspsychologie und ihre Anwendungen, pp. 59–71. Kohlhammer, Stuttgart (2005)
15. Rheinberg, F.: Motivationsdiagnostik, 7th edn. Kohlhammer Verlag, Stuttgart (2008)
16. Csikszentmihalyi, M.: Das Flow-Erlebnis: Jenseits von Angst und Langeweile: Im Tun aufgehen, 8th edn. Klett-Cotta Verlag, Stuttgart (2000)
17. Csikszentmihalyi, M.: Ein theoretisches Modell des optimalen Erlebens: Einführung. In: Csikszentmihalyi, M., Csikszentmihalyi, I.S. (eds.) Die außergewöhnliche Erfahrung im Alltag - Die Psychologie des Flow-Erlebnisses, pp. 15–27. Klett-Cotta Verlag, Stuttgart (1995)
18. Rheinberg, F., Vollmeyer, R.: Flow-Erleben in einem Computerspiel unter experimentell variierten Bedingungen. Zeitschrift für Psychologie **15**, 161–170 (2003)
19. Rheinberg, F., Manig, Y., Kliegl, R., Engeser, S., Vollmeyer, R.: Flow bei der Arbeit, doch Glück in der Freizeit: Zielausrichtung, Flow und Glücksgefühle. Zeitschrift für Arbeits- u. Organisationspsychologie **51**, 105–115 (2007)
20. Rheinberg, F., Vollmeyer, R., Engeser, S.: Die Erfassung des Flow-Erlebens. In: Stiensmeier-Pelster, J., Rheinberg, F. (eds.) Diagnostik von Motivation und Selbstkonzept, pp. 261–280. Hofgrefe, Göttingen Bern Toronto Seattle (2003)
21. Consolvo, S., Walker, M.: Using the experience sampling method to evaluate ubicomp applications. IEEE Pervasive Comput. **2**, 24–31 (2003)
22. Larson, R., Csikszentmihalyi, M.: The experience sampling method. In: Csikszentmihalyi, M. (ed.) Flow and the Foundations of Positive Psychology, pp. 21–34. Springer, Dordrecht (2014)
23. Rheinberg, F.: Zweck und Tätigkeit. Hogrefe, Göttingen (1989)
24. Rheinberg, F., Iser, I., Pfauser, S.: Freude am Tun und/oder zweckorientiertes Schaffen? Zur transsituativen Konsistenz und konvergenten Validität der Anreizfokus-Skala. Diagnostica **2**, 174–191 (1997)
25. Camerer, C., Hogarth, R.: The effects of financial incentives in experiments: a review and capital-labor-production framework. J. Risk Uncertainty **1–3**, 7–42 (1999)
26. Crawley, J.: The R Book, 2nd edn. Imperial College London at Silwood Park, London (2013)
27. Abeler, J., Falk, A., Götte, L., Huffman, D.: Reference points and effort provision. Am. Econ. Rev. **101**, 470–492 (2011)
28. Engeser, S., Rheinberg, F.: Flow, performance and moderators of challenge-skill balance. Motiv. Emot. **32**, 158–172 (2008)
29. Robson, K., Plangger, K., Kietzmann, J., McCarthy, I., Pitt, L.: Game on: engaging customers and employees. Bus. Horiz. **59**, 29–36 (2016)
30. Hamari, J., Koivisto, J., Sarsa, H.: Does gamification work? - A literature review of empirical studies on gamification. In: 47th Hawaii International Conference on Gamification System Sciences (HICSS), pp. 3025–3034 (2014)
31. Falk, A., Heckman, J.: Lab experiments are a major source of knowledge in the social sciences. Science **326**, 535–538 (2009)
32. Aparicio, A., Vela, F., Sánchez, J., Montes, J.: Analysis and application of gamification. In: Proceedings of the 13th International Conference on Interacción Persona-Ordenador, 17 (2012)

Students' Choices

A Comparative Study of a Gamified and a Non-gamified Question-Based Learning App in Graduate Education

Heinrich Söbke(✉) and Laura Weitze

Bauhaus-Institute for Infrastructure Solutions (b.is),
Bauhaus-Universität Weimar, Weimar, Germany
{heinrich.soebke, laura.weitze}@uni-weimar.de

Abstract. Serious Games are applied to foster motivation in accomplishing real-world tasks. However, their efficacy can be considered as being dependent on personal preferences of the recipients. In this exploratory study, two multiple choice question-based apps have been evaluated as a learning tool in a university course of an environmental engineering study program: a commercial quiz app (*QuizUp*) and a professional learning app (*Skive*). At the beginning of the course each student (N = 15) chose one of these apps as a mandatory learning aid for factual knowledge over the complete course period. The study results, collected by learning diaries, a questionnaire and results of the final test, confirm the type dependence of learning experiences created by these tools. Further, it acknowledges mostly the maturity of the used question-based learning app, but reveals current limitations of both, the quiz app and the didactical context.

Keywords: Quiz app · Mobile learning · Learning app
Multiple choice question · *Skive* · *QuizUp*

1 Introduction

Repeated answering of multiple choice questions (MCQ) leads to learning due to the so-called "testing effect" [1, 2]. The recent availability of mobile devices and mobile internet creates an almost ubiquitous accessibility of those questions and fosters casual mobile learning opportunities [3, 4]. Even some quiz apps are opened to user defined content. Thus they can be seen as potential learning tools [5]. In a previous study about the usage of the commercial quiz app *QuizUp* [6] in a university master course the authors found that there can be differentiated roughly between *Gamers* and *Learners* [7]: *Gamers* are open to the appeal and mechanics of quiz apps and use them intensively. In contrast, *Learners* focus on reaching their learning goals and perceive the game mechanics as time-wasting burden. Fortunately, there exist already mobile learning apps that are based on MCQs, too. Therefore, the contents between quiz and learning apps are easily transferable. This led to a study design, which projected to use both genres of apps in parallel. Students were offered the possibility to choose their preferred app genre. The main research question of this exploratory study is observing the effects to learning and acceptance, if learners have a choice. After having

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 105–114, 2017.
https://doi.org/10.1007/978-3-319-71940-5_10

experienced partly disaffirmation of gamified quiz apps by students in previous studies [7, 8], this study aims at investigating if a gamification layer is perceived as beneficial by students or if they prefer a straight learning approach by a dedicated learning app. Further, the parallel and open usage of both tools should lead to reflection of students about advantages and disadvantages of each app genre.

2 Study Design and Methodology

Skive [9] and *QuizUp* [6] have been chosen as tools. Both apps use identical MCQs as learning content. The gamification layer of *QuizUp* is the main difference. *Skive* is a question-based learning tool explicitly designed for university learners. It offers an MCQ feature (see Fig. 1). Contents are structured by institutions and their courses. Conversely, *QuizUp* is a commercial quiz app, which can be equipped with user-generated content [5]. Questions are answered in *QuizUp* in so-called matches, i.e. seven questions are operated in a competition against an opponent. This opponent can be either another player or an app-created bot. The maximum time to answer a question is restricted to ten seconds (see Fig. 1). In contrast, *Skive* allows answering categorized questions sequentially without any time limit. Both tools offer a history of answered questions that includes the correct answers.

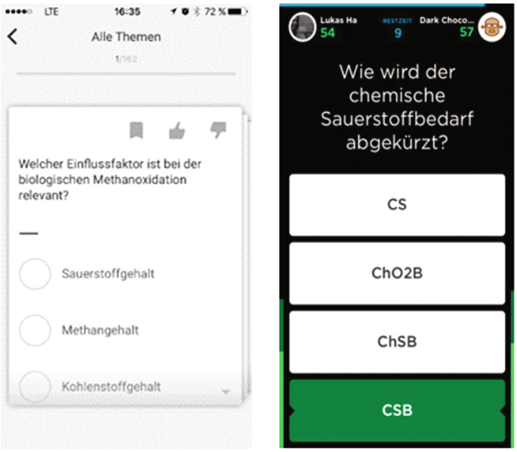


Fig. 1. Screenshots of Skive (left) and QuizUp (right) with sample questions in German

The study has been carried out within the course "*Technical Air Pollution Control*" in the master study program "*Environmental Engineering*" in the fall semester 2016. Students chose their preferred learning tool, either *Skive* or *QuizUp,* as a mandatory learning tool during the registration process. Only three participants of the altogether fifteen participants (i.e., 20% of N = 15) have replied with an explicitly explained (max. 60 words) choice: two voted for *Skive* (main reason was that they wanted to get familiar with another tool, as they already knew *QuizUp* from a previous application [7]), one voted for *QuizUp*. The other students have been assigned randomly an app, ensuring that both factions have an almost equal size. Throughout the semester, students were required to provide proofs of a minimum activity at three distinct points of time. They handed in screenshots of the app demonstrating that each question has been answered at least once. Furthermore, students have been instructed to write a learning diary in a semi-structured form (questionnaire items and free text questions about their experiences and preferences) every two weeks, altogether four times. Standardized items of the *Game Experience Questionnaire* (GEQ) [10] and the *Fragebogen User Experience* (FUX) [11] have been included. Finally, a concluding questionnaire (49 questions) has been issued.

MCQs have been developed by students throughout the course: each student had to create one question per week. As a preparation a guideline for creating good MCQs was provided to students [12]. After formal reviews, finally 161 questions have been made available to students (*Skive*-course: "*Biologische Verfahren der Luftreinhaltung*" of the institution "*Bauhaus-Universität Weimar*" and *QuizUp*-topic "*Luftikusse*"). Ten of those questions have been included in the final test.

3 Results

A selection of the results is presented in the sequence of their gathering: **learning diaries** were written during the semester and the **questionnaire** has been filled ahead of the **final test**. Textual comments in the learning diaries are summarized together with the results of the questionnaire.

3.1 Learning Diaries

Main part of the learning diaries, which have been filled out at four distinct point of times (here called *Timebases* (TB)), was the inline variant of the IJsselsteijn et al.'s *Game Experience Questionnaire* (GEQ), which consists of fourteen items for determination of seven variables [10]. First motivation for this approach was to check if there are differences between both apps regarding the perception as a game. Further, a potential development over time of measured variables should be examined. Figure 2 shows the comparison between *Skive* and *QuizUp*. There are partly only slight differences. However, they seem to be reasonable. First, there is almost no difference for *Competence* and *Challenge*. This may be related to the similar tasks (namely, answering MCQs), which have to be accomplished. *QuizUp* leads in values for *Sensory + Imaginative Immersion* and *Flow*. A reason might be the app's game characteristics. In contrast, *Skive* shows better values of *Positive Affect*, *Negative Affect* and *Tension*. This could be traced back in parts to the facts of "mandatory" game play and disturbing and time demanding advertisements in *QuizUp*.

The temporal development of variables is illustrated in Fig. 3. Besides at least one constant variable there can be assumed some regular developments. There seems to be almost no change for the variable *Challenge*. Values for *Tension* are growing except for TB 3 and TB 4. In general, the values for TB 3 seem to be influenced by a specific, extraordinary event (e.g. complex or unclear questions, especially complex learning content or extra effort for other courses): *Competence*, *Immersion*, *Flow* and *Positive Affect* are reduced whereas *Tension* and *Negative Affect* appeared increased compared roughly to TB 2. In general, further studies could either reject or approve here suggested development patterns like a decreasing flow.

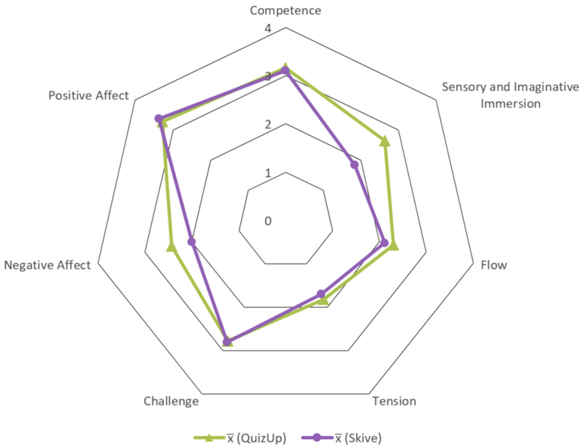


Fig. 2. GEQ: Comparison of *Skive* and *QuizUp* (N = 15)

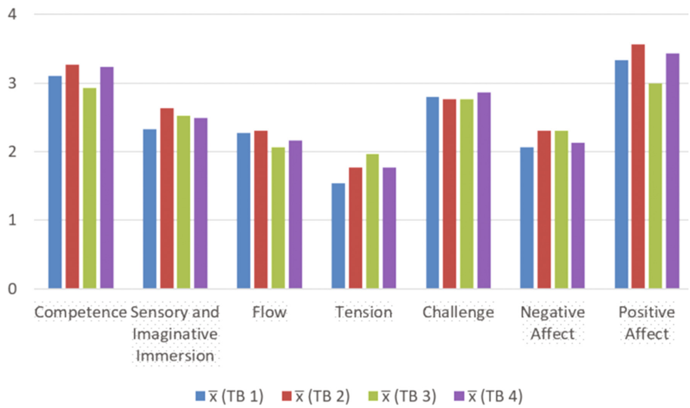


Fig. 3. GEQ: Development of variables over time (N = 15)

3.2 Questionnaire

The questionnaire was answered by all fifteen participants and contained altogether forty-nine questions in the categories *Experiences and Attitudes* (3 questions), *Creation of Questions* (4), *Tool Usage* (7), *Skive* (5), *QuizUp* (7), *Usability* (4), *Didactical Context* (13), *Demography* (2) and *General Remarks* (2). 33% of the participants were female. Years of birth ranged from 1989 to 1993. Tool-specific questions had to be answered only by those students who used the tool (*Skive*: N = 8, *QuizUp*: N = 7). Twenty five questions used Likert scales to ask for a level of agreement with a statement. These questions comprised altogether 175 items. In the following noticeable results are discussed. The questionnaire was issued in German (according to the course language); the given translations have been made for this report.

Experiences and Attitudes. Participants were asked about their **game-play frequency** in general. Based on a 5-point Likert scale (1: never – 5 very often), mobile phone games reached the highest value ($\overline{X} = 2.8$), whereas learning games had the lowest value (1.5). Board games (2.5), multiplayer online games (2.6) and quiz games (2.6) were genres that received average popularity values. The next question should capture an attitude towards **game-based learning** (GBL). On a 5-point scale (1: false, 5: true), the statement "I already have played a learning game successfully" reached an average value of 2.5. In contrast, the claim "That does not work" reached only 1.9 points. Participants value GBL as a "good opportunity to deal with a topic" in a self-directed manner. However, the by far strongest approval got the assertion "I prefer learning in a more direct manner" (3.9). Another question asked for the attitude towards question-based apps in learning contexts. Most striking values were received for the statement "I use such apps only if their contents are test-relevant." (3.8 out of 5 points) and "Ranking lists and competitions are not meaningful to me" (3.7). A further question examines the **attitude** to **the technical domain** of the study course and the program. The items revealed a positive attitude and a high identification with the discipline "Environmental Engineering is a relevant discipline" (4.1), "I would like to work in the field of Air Pollution Control" (3.5) and "I have chosen this course very deliberately." (4.2). In contrast, the statements "I do not like the subject of *Air Pollution Control*" (1.7) and "I just study to receive the leaving certificate" (1.3) have been rejected. Altogether, these figures draw a picture of motivated, game-aware, but not intensively game-playing participants, who are at the same time open and sceptical towards GBL and educational technology.

Tool Usage. 60% of students indicated that they have had at least a look at the other tool. Thus, they are able to compare the apps. Two third (ten out of fifteen) have accessed the apps via smartphone, although for both apps there exists also a web browser version. Therefore, this kind of usage can be considered as kind of **mobile learning**. Furthermore, it is interesting if there happens **casual or dedicated usage**. Figure 4 shows the results. In general, participants were forced to use the apps by deadlines. However casual usage contexts show the flexibility of these mobile learning apps (e.g. during travelling, during waiting times and even on the toilet). The **extrinsic motivational effect of deadlines** is underlined by a question for the triggers for learning sessions. The statements, that usage is integrated into daily routines is rejected ($\overline{X} = 1.5$ on a 5-point scale, 1: false – 5: true). Contrasting, "Only before deadlines" reaches a consent of 3.9. Noteworthy is the negation of common learning sessions with fellow students (1.4). There seems to be almost no social dimension of the app usage.

***Skive* and *QuizUp*.** Besides the immediate goal for students to demonstrate that all questions have been answered correctly at certain deadlines, both apps offer further features. Usage of and attitude towards these features has been investigated in this category of questions. Although both tools offer a broad range of topics, only four participants (27%) had a look at other topics. Further features of *Skive* (e.g. flip card, bookmarking questions, commenting questions, sorting and filtering of questions) have been used only marginally. There was the wish to immediate further investigation on unclear or imprecise questions (3.5 (*Skive*) resp. 3.9 (*QuizUp*) on a 5-point scale, 5: true). Desire for **competition** (1.8) and **social networking** with fellow students (1.9 (*Skive*) resp. 1.7 (*QuizUp*)) has been **rejected**. Consequently, it was negated that the app led to new acquaintances (1.1). However, the statement that **competing with friends** has a considerable impact on motivation (3.9) has been approved by *QuizUp* players. Another motivational factor has been **seeing one's own questions** in the tool (3.4 (*Skive*) resp. 3.6 (*QuizUp*)), **Main activity** during *QuizUp*-usage has been answering questions (4.9 on a 5-point scale: 1: never – 5: during each session). The only two other activities, which got an average value of more than two are "looking at ranking lists" (2.9) and "reviewing the history of matches" (2.4). While the first activity may be an evidence for the desire for progress-approval and to compare one's activities with those of other course members, the latter one points to reflection processes, which are hardly possible during a match due to limited answering times. **Advertisements** in *QuizUp* have been marked as **disturbing** (4.0) and interrupting the game-flow (3.6). A finding from a previous study [8] was that entertainment topics receive a higher **acceptance** than educational topics. Besides the fact that the textual complexity of educational questions is higher compared to those of entertainment questions [13], this study brings up other potential explanations. The quality of educational questions seems to be worse (1.2 of a 5-point scale; 1: true – 5: false), the pressure to learn spoils a positive experience (1.8) and participants do not like quiz apps in general (1.8). However, in another question responders conceded that having fun does not depend on the educational or entertainment purpose of the topic, but on one's personal interest in the topic (4.1 of a 5-point scale with 1: I do not agree – 5: I agree completely).

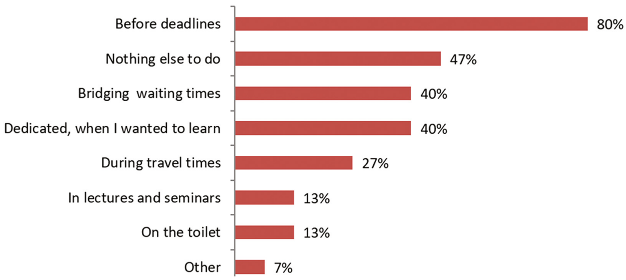


Fig. 4. Usage contexts of the app (N = 15, multiple selections possible)

Didactical Context and General Remarks. A first question should reveal basic approaches and technologies, which are used by participants for learning. As shown in Fig. 5, students are aware that there is a need to learn ("I do not have to learn", 2.2) and they practice dedicated learning before exams (4.5). Mostly they use self-created lecture scripts excerpts (4.3) or paper flip-cards (3.4). Often, they participate in peer-groups (3.4). In contrast, they use digital learning tools seldom (1.7). Besides these given options, free text answers pointed to other media: *short internet video clips*, creation of, and listening to *audio memos*. Students estimated that 43% of the questions are relevant for their future professional work and 66% for the final test. Fourteen of fifteen participants (93%) approved that the tools caused learning effects. Especially, they named numbers and terms, which are kinds of factual knowledge. A further by-effect of student-generated questions is the formation of technical focusses, which in turn have led for some students to specific learning engagement. Asked for their choice in a future course, thirteen of fifteen students opted for *Skive*. As reasons for this unambiguous vote were named features of *Skive* (deliberate repetition of problem question, status-indicators of work-progress), but at the same time weaknesses of *QuizUp*: disturbing waiting times, time limits for answers, no possibility to repeat specific questions, no learning algorithms).

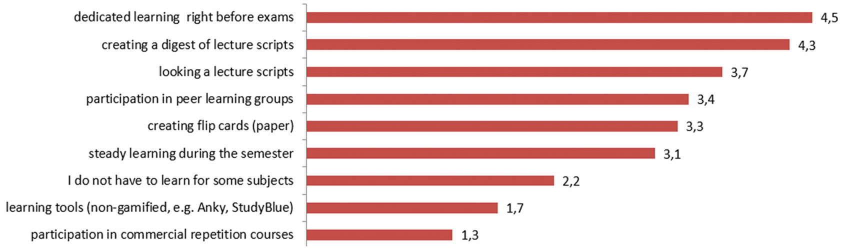


Fig. 5. Approaches and technologies for learning (N = 15, 5-point scale, 1: false, 5: true)

Summarizing remarks drew a mixed picture. On the one hand, positive aspects were mentioned, like variety in daily learning routines. On the other hand, the reason to exist of such tools has been questioned at all: "self-created script excerpts are totally sufficient". Further, these apps were classified as "not appropriate for a university educational institution and their educational goal of self-directed learning".

Creation of Questions and Usability. These categories are not in the primary focus of this report. However, there are outcomes that should be mentioned here. The statements "Questions should be verified by teaching staff" (4.1), "Some questions were unclear" (3.6) and "It was apparent that questions have been created by students" (3.4) got high approval rates. Interestingly, participants indicated that they have created their own questions with a special focus on quality (4.0). FUX ("Fragebogen User Experience"), a set of standardized items for measuring usability [11], has been applied. Most salient results are ratings as "interesting" (6.9 on a 10-point scale) and "precious" (6.4). Both apps have been described as "easy to understand", "easy to learn" and "to operate without further instructions" (each 4.3 points on a 6-point scale).

3.3 Final Test

The final test was taken by thirteen participants. Five of them used previously *QuizUp* (the remaining two missed the test for the reason of illness); the other eight were *Skive* users. The test contained - besides other tasks - ten questions strongly connected to those questions, which were previously trained through the question-based apps. Answering them correctly resulted in 10 points. Altogether, examinees could reach 60 points. Table 1 shows the results, which cannot be considered as quantitatively representative due to the small number of participants. *Skive* users reached better results in the question section: 71% instead of 61% with a lower standard deviation (SD). For the complete test, the advantage for *Skive* users decreased to 3%. A possible reason for this reduction could be that previously trained questions about factual knowledge are less complex than calculation tasks that require knowledge about systems. The functional advantage of *Skive* could lead to a greater acceptance and usage and in turn to better memorization results.

Table 1. Results of the final test (Percentage of achieved points)

	Part questions		Complete test	
	$\bar{x}$	SD	$\bar{x}$	SD
QuizUp-users (N = 5)	61%	20%	60%	16%
Skive-users (N = 8)	71%	16%	63%	22%
All (N = 13)	66%	21%	62%	20

4 Discussion

The apparent result of this study is that the learning tool *Skive* is preferred by far to the quiz app *QuizUp*. Results of the final test suggest a potential better learning outcome of *Skive*. However, the functionality and user experience of *QuizUp* was impacted by a few shortcomings. Especially four issues have to be taken into account. First, a careful design of a social context is necessary, as competition with friends is a major motivational elements when it comes to playing quizzes [14]. In formal contexts of education, it seems that social interactions have to be explicitly spurred by instructional design and do not evolve spontaneously. Thus, as in the study matches mostly were played against bots, an important gain of games was not developed here. Second, intrusive advertisements impeded the gaming and learning experience seriously. Different kinds of advertisements (ad bar at the bottom of the screen, not instantly interruptible ads between matches) have been introduced during the study. As *QuizUp* did not offer a pay-release, there was no opportunity to avoid these advertisements[1]. Third, players had no influence on the asked question. In the context of a learning tool, there should be an automated preference of presenting those questions, which have not

[1] In the meantime, there is a paid (approx. € 3) ad-free app variant available.

been asked or answered correctly, yet. Finally, a fourth restriction is given by the limited answering time. This time limit maybe beneficial as a game mechanic. However, when players are in a learning process, they often remain unsatisfied. A potential trade-off could be a question specific time span, which is increased for complex questions.

In general, the necessity for such question-based tools was questioned by parts of the audience. However, the appropriateness of such a tool depends on the learning content. Other participants stated that the questions covered suitable kinds of knowledge. Further, for example, in the field of medicine comprehensive exams (e.g. "Physikum") are based on MCQs and their preparation is assisted by digital learning apps [15].

Moreover, the quality of student-generated MCQs led to irritation among the audience. In general, question design is a demanding task and even if guidelines are applied, there is the need of quality assuring measures. They can be performed either by fellow students (e.g. collaborative question creation [16, 17]) or by student-instructor feedback cycles. The first alternative demands technical requirements, the latter one does not scale and shifts considerable workload to the teaching staff.

Of course, this study's sample size is very small to deliver quantitatively reliable results. So, the found results – except for the discussed obvious outcomes regarding *QuizUp* – should be considered as hypotheses (resp. as hints to hypotheses), which still have to be validated.

5 Conclusions

In this study, thirteen of fourteen students opted for *Skive* as a learning tool in subsequent courses. This seems to be an unambiguous result for the employment of a pure learning app in higher education learning contexts. In contrast to quiz apps, these apps can be considered as having reached application maturity. However, there could be identified important shortcomings of the used quiz app and the design of the application context. Their correction could bring quiz apps to become a further learning tool which develops its strengths in well-designed learning contexts, e.g. by reviving common learning sessions [7]. Future work needs to identify such learning scenarios in order to enable quiz apps to become an accepted learning tool and contribute to didactically diversified courses. For both variants, a rigid quality management for the content is necessary if questions are created by learners.

References

1. Karpicke, J.D., Roediger, H.L.: The critical importance of retrieval for learning. Science **319**, 966–968 (2008)
2. Agarwal, P.K., Karpicke, J.D., Kang, S.H.K., Roediger, H.L., McDermott, K.B.: Examining the testing effect with open- and closed-book tests. Appl. Cogn. Psychol. **22**, 861–876 (2008)
3. Gikas, J., Grant, M.M.: Mobile computing devices in higher education: Student perspectives on learning with cellphones, smartphones & social media. Internet High. Educ. **19**, 18–26 (2013)

4. Wu, W.H., Wu, Y.C.J., Chen, C.Y., Kao, H.Y., Lin, C.H., Huang, S.H.: Review of trends from mobile learning studies: a meta-analysis. Comput. Educ. **59**, 817–827 (2012)
5. Woods, B.: QuizUp launches tools for creating your own trivia categories and questions. http://thenextweb.com/apps/2015/09/24/quizup-launches-tools-for-creating-your-own-trivia-categories-and-questions/#gref
6. Plain Vanilla: QuizUp - Connecting people through shared interests, https://www.quizup.com/
7. Weitze, L., Söbke, H.: Quizzing to become an engineer - a commercial quiz app in higher education. In: Pixel (ed.) Conference Proceeding, New Perspectives in Scienze Education, 5th edn., Florence, pp. 225–230. libreriauniversitaria it Edizioni (2016)
8. Söbke, H., Weitze, L.: The challenge to nurture challenge students' perception of a commercial quiz app as a learning tool. In: Wallner, G., Kriglstein, S., Hlavacs, H., Malaka, R., Lugmayr, A., Yang, H.-S. (eds.) ICEC 2016. LNCS, vol. 9926, pp. 15–23. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46100-7_2
9. qLearning Applications GmbH: Skive (2016)
10. IJsselsteijn, W.A., De Kort, Y.A.W., Poels, K.: The Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games (2013). Manuscript in Preparation
11. Müller, J.: FUX (Fragebogen User Experience) (2012)
12. Söbke, H., Kämmerer, F.: Vermessene Fragen: Metriken als Ansatz automatisierter analytischer und konstruktiver Qualitätssicherung von Mehrfachauswahlfragen für mobile digitale Medien. In: Pfau, W., et al. (eds.) Teaching Trends 2016-Vielfalt in der Lehre, Clausthal-Zellerfeld, 10/11 November, pp. 153–162. Waxmann, Münster (2016)
13. Kämmerer, F., Söbke, H., Hartung, L.: Schnell zu erfassen: Ein Komplexitätsmaß für Mehrfachauswahlfragen. In: Igel, C., et al. (eds.) Bildungsräume DeLFI 2017, 5. bis 8 September 2017, Chemnitz Proceedings, pp. 155–166. Gesellschaft für Informatik, Bonn (2017)
14. Söbke, H.: Space for seriousness? Player behavior and motivation in quiz apps. In: Chorianopoulos, K., Divitini, M., Hauge, J.B., Jaccheri, L., Malaka, R. (eds.) ICEC 2015. LNCS, vol. 9353, pp. 482–489. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24589-8_44
15. Miamed GmbH: AMBOSS. https://www.miamed.de/amboss
16. Parker, R., Manuguerra, M., Schaefer, B.: The Reading Game – encouraging learners to become question- makers rather than question-takers by getting feedback, making friends and having fun. In: Carter, H., Gosper, M., Hedberg, J. (eds.) 30th ascilite Conference 2013 Proceedings, pp. 681–684. Macquarie University (2013)
17. McClean, S.: Implementing PeerWise to engage students in collaborative learning. Perspect. Pedagog. Pract. **6**, 89–96 (2015)

Towards Implementing Gamification in MOOCs

Alessandra Antonaci(✉), Roland Klemke, Christian M. Stracke, and Marcus Specht

Research Centre for Learning, Teaching and Technology, Welten Institute, Open University of the Netherlands, P.O. Box 2960, 6401 DL Heerlen, The Netherlands
{alessandra.antonaci,roland.klemke,christian.stracke, marcus.specht}@ou.nl

Abstract. Gamification is well known as a design strategy used to generate a change in users' behaviour, such as motivation. However, while in recent years interest in it has been growing, empirical evidence on the effects that the application of game elements can generate on users' behaviour is still lacking. We present the results of a study as a step towards designing gamification with better understanding of the possible effects that each game element could generate on end users. By involving three groups of experts: game designers, learning scientists and specialists in technology-enhanced learning (TEL), we assessed a selected number of 21 game design patterns in relation to the effects these could generate on learning performance, goal achievement and engagement of learners if implemented in a Massive Online Open Course (MOOC). Based on quantitative and qualitative data collected, 9 game elements have been selected to be further investigated.

1 Introduction

In the last 10 years, gamification has been applied in several fields, such as: marketing, trading, training, fitness, software engineering, etc. The highest number of applications can be found in education [1, 2]. "Education is a fundamental human right and essential for the exercise of all other human rights. It promotes individual freedom and empowerment and yields important development benefits" [3]. With the aim of bringing education to all, phenomena such as Open Education and in particular MOOCs have started to grow in recent years.

MOOCs, taking advantage of the Internet, could hypothetically bring knowledge to everyone and for free. Despite this potential, MOOC studies have highlighted several limits, such as the lack of users' engagement and their high dropout rate [4]. To overcome these gaps our idea is to apply gamification in MOOC to enhance users' goal achievement, engagement and also impact on users' learning performance.

Conceptual examples of gamification applied in MOOC environments can be found as detailed in [4]. A first attempt at raising awareness on gamification by empirical evidence can be found in [5], where 40 suitable game mechanics to engage students in MOOCs were identified and rated by 5,020 participants through an online survey for

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 115–125, 2017.
https://doi.org/10.1007/978-3-319-71940-5_11

their level of engagement. Among the 40 selected, 10 game mechanics were evaluated with the highest level of engagement (virtual goods; three different types of points; leader boards; trophies and badges; peer grading and emoticon feedback; two types of games) by users. The engagement level of these mechanics was defined based on users' self-perception, not on an empirical basis and space was not given to the designers' point of view. Which effects did the designers aim to stimulate in users? Did the perception of users match with the intention of the designers? With the purpose of testing this relationship between designers' intention and effects on users' behaviour and perception, in this paper we present our first step towards implementing gamification in MOOCs by involving three groups of experts that are invited to evaluate a selection of game design patterns[1] (GDPs) for their suitability in a MOOC application.

The three groups of experts come from different but complementary fields: (1) game design, from them we expected the ability to evaluate the effects of specific GDPs in a given scenario from a game perspective; (2) learning science, who could judge the GDPs from a didactic and educational perspective; and (3) the TEL field who can evaluate both perspectives and rate applicability and feasibility of the GDPs chosen. Each group of experts assess each of the 21 available GDPs from their perspective in accordance with the scenario of application (MOOC) and one of the following purposes of intervention: p1- enhancing MOOC users' learning performance via gamification; p2- enhancing MOOC users' goal achievement and p3- enhancing MOOC users' engagement.

The 21 GDPs chosen come from Björk and Holopainen's collections [6]. These were selected in accordance with the following criteria (1) the frequent use of a GDP in the literature, (2) the applicability of a GDP in a multi-user environment, and (3) our hypothesised impact of the selected patterns on learners' engagement, goal achievement, or learning performance. As a result, the following GDPs were presented to 42 experts to be validated:

1. *Avatars/Characters*, that are "abstract representations of the person".
2. *Time Limit*, it can be related to action completion or goal achievement.
3. *Levels*: "a level is a part of the game in which all players' actions take place until a certain goal has been reached or an end condition has been fulfilled".
4. *Communication Channels*, "are the medium and the methods players can use to send messages to other players".
5. *High Score Lists*, display players' rank and give them the chance to compare themselves against other players "who have previously played the game".
6. *Score,* "is the numerical representation of the player's success in the game".
7. *Status Indicators*, via them users can gain information about "a certain part of the game state or other players".

[1] In this paper the term game elements, game mechanics and game design patterns are used as synonyms even if the authors are aware of their differences. In particular, game elements are attributes of a game (such as score), while game mechanics are "any part of the rule system of a game that covers one, and only one, possible kind of interaction that takes place during the game" [8]. As for Game Design Pattern, (GDP), "The origin of the concept of "design patterns" hails from the field of architecture and in particular was coined by Christopher Alexander" [9]. Applying GDP is also a method of codifying game design knowledge in separate but interrelated parts.

8. *Public Information*, is information of the game state "available during the game to people other than the players".
9. *Storytelling*, the act of telling stories within the game.
10. *Rewards*, the player receives something for completing goals or levels in the game.
11. *Goal Indicators*, give information to players about their current goals in the game.
12. *Stimulated Planning*, enables users to plan certain aspects or actions in the game.
13. *Clues*, "are game elements that give the players information about how the goals of the game can be reached" [6].
14. *Cooperation*, "i.e. coordinate their actions and share resources, in order to reach goals or sub-goals of the game".
15. *Limited Planning Ability*, implies that players do not have all the information.
16. *Competition*, "is the struggle between players or against the game system to achieve a certain goal".
17. *Team Play,* "Players in a group or a team coordinate their actions, abilities, and roles in order to reach a common goal".
18. *Replayability*, enables a re-doing of the game or level.
19. *Smooth Learning Curves*, enable the chance to "smoothly progress from novice to master".
20. *Handicaps*, enable the chance to provide an easier gameplay in order to make all players have the same chance of succeeding.
21. *Empowerment*, "Players feel that they can affect the events and the final outcome of a game" [6].

2 Study

Procedures. Participants were introduced to "MOOCs" and "Gamification". The game designers were invited to take part in the focus group as part of a game design workshop and were divided into six groups assigned to the three intervention purposes p1–p3 (two groups for each purpose). These three purposes were identified in accordance with a literature review study related to MOOCs gaps. They are interdependent from each other and were settle to facilitate the participants' conceptualization. The topic of the MOOC was predefined as cyber-security. Each group elaborated a concept that was presented to the other colleagues. All participants filled out the survey, comprising 2 questions for each of the 21 GDPs selected: a closed question, rating the GDPs in relation to the purpose selected (p1–p3) using a scale from 0 ("strongly negative effect") to 4 ("strongly positive effect"). The second question for each GDP was optional and open; here participants could detail the advantages and/or disadvantages of using the given GDP for the specific purpose.

Methods. Two methods were used to assess the GDPs selected for the gamification design of a MOOC: a survey and a focus group. The survey was designed with the aim of validating our GDP selection and collecting feedback from our target population. Using a scale from 0 (strongly negative effects) to 4 (strongly positive effects), participants were asked to rate a total of 21 game elements. The game elements proposed to our audience population were rated in accordance with a specific scenario of application and a given purpose. The application scenario was a MOOC while the purpose of the intervention could be chosen from the following options: Enhancing users' learning performance via gamification (p1); Enhancing users' goal achievement via gamification (p2); Enhancing users' engagement via gamification (p3). Secondly, the focus group was designed and conducted for game designers with the aim of inviting them to conceptualise a gamified MOOC using the game elements that they deemed most relevant for one of the specific MOOC purposes given (p1–p3). The subjects could independently decide on which of the three gamification purposes of the intervention they wanted to focus (p1–p3).

Participants. A total of 42 subjects took part in our study, they come from different backgrounds and fields: 17 game design experts; 9 learning scientists and 16 TEL experts. Table 1 presents an overview of participant personal details and distribution among the purposes chosen.

Table 1. Sample distribution and details

	Game designers	Learning scientists	TEL experts
Age range per group	18–24 years	25–34 years	25–34 years
P1- enhancing MOOC users' learning performance	6 (out of 17)	4 (out of 9)	5 (out of 16)
P2- enhancing MOOC users' goal achievement	6 (out of 17)	3 (out of 9)	4 (out of 16)
P3- enhancing MOOC users' engagement	5 (out of 17)	2 (out of 9)	7 (out of 16)

The level of awareness of the two concepts: "gamification" and "MOOC" was also investigated and among the game designers 2 of them had followed one MOOC previously and they had all heard about gamification before. For the learning scientists, one of them had designed a MOOC, 3 (out of 9) had participated in one MOOC before and 1 (out of 9) in more than one MOOC, one of them had heard about gamification before. As for the TEL experts 2 (out of 16) had designed a MOOC before, 4 (out of 16) had followed one MOOC previously, 10 (out of 16) had followed more than one MOOC and 2 (out of 16) none. All had "heard about gamification before".

3 Results

3.1 Quantitative Data

In this section, we present the data collected with the questionnaire divided into three sub-sections, according to the participant audience. Starting with the **Game design experts' evaluation** Fig. 1 represents a summary of how the experts in game design rated the selected GDPs in accordance with the effects they expected related to the three purposes identified (p1–p3) of our gamified intervention within a MOOC.

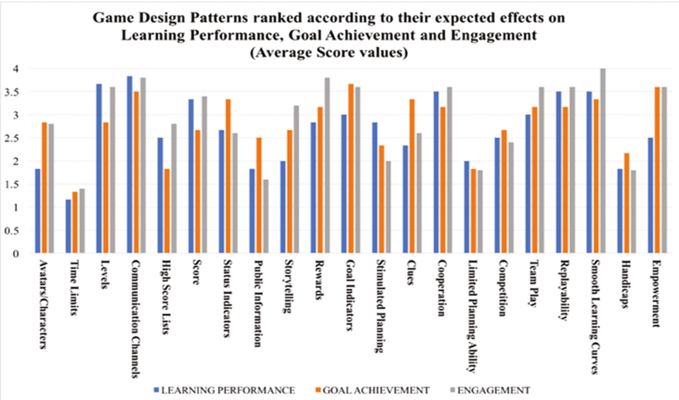


Fig. 1. Representation of the GDP average values rated by game designers according to their expected effects on p1 (blue), p2 (orange) and p3 (grey) within a MOOC. (Color figure online)

More in detail, with respect to p1, the GDPs most ranked are: *Communication Channels* with average score of 3.83; *Levels* average 3.67, and with equal score and average 3.5 *Cooperation*, *Replayability* and *Smooth Learning Curves*. The lowest score is *Time Limits* with a mean value of 1.5, which considering the scale used allocates this GDP under the "neutral effect" (2) closer to the range "negative effect" (1).

With respect to p2 (goal achievement), the highest ranked GDPs by the game designers are: *Goal Indicators* with an average score of 3.67, *Empowerment*, average 3.6, and *Communication Channels*, average 3.5. It needs to be underlined that none of the top 10 GDPs with high mean values was scored with 0 (strongly negative effects).

Rating GDP effects on p3 (engagement), the GDPs with the highest average scores received are: *Smooth Learning Curve*s with an average of 4, *Communication Channels* and *Reward* both with an average score of 3.8 and as for the other two purposes detailed, *Time Limits* registered the lowest average score of 1.4.

The learning scientists' evaluation. Among the learning scientists, the highest number (4 out of 9) chose to express their point of view on p1(learning performance), while the rest evaluated p2 (goal achievement) (3 out of 9) and p3 (engagement) (2 out of 9). With respect to p1, the learning scientists assigned the highest score to the following GDPs: Levels with average score of 4; *Empowerment*, average 3.75, and with equal score and average 3.5 *Avatar/Characters*, *Storytelling* and *Clues*. For p2, the highest ranked GDPs are *Smooth Learning Curves*, average score 4 and the equally scored *Clues* and *Empowermen*t, average 3.67. Looking at p3, the learning scientists evaluate the following GDPs with an average score of 4: *Storytelling*, *Clues* and *Empowerment* (Fig. 2).

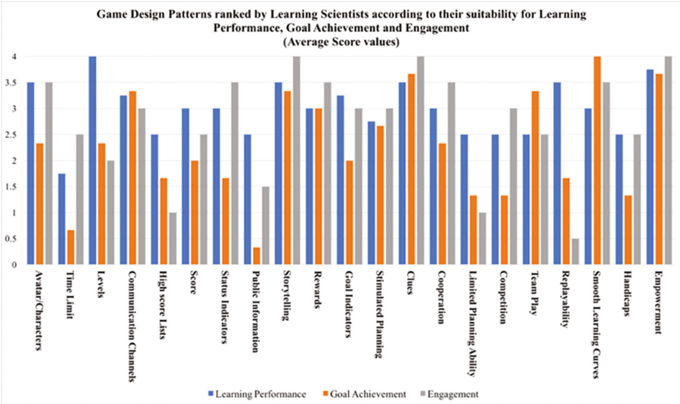


Fig. 2. Representation of the GDP average values rated by learning scientists related to p1 (blue), p2 (orange) and p3 (grey) (Color figure online)

The TEL experts' evaluation. As shown in Fig. 3, the TEL experts (5 out of 16) evaluated the following GDP effects on p1 with the highest score: *Levels* and *Smooth Learning Curves* both with an average score of 3, and with equal score and average of 2.8: *Storytelling*, *Replayability* and *Empowerment*. Four participants selected p2 and according to their evaluation, the GDPs with highest scores were: *Goal Indicators* with an average score of 4; *Levels*, average 3.5, and, with equal score *Replayability* and *Smooth Learning Curve* with an average of 3.25. With respect to gamifying a MOOC to enhance users' engagement (p3), the GDPs that collected the highest scores are *Communication Channels, Score*, *Goal Indicators*, *Cooperation* and *Smooth Learning Curve*, all with the same average score of 3.43.

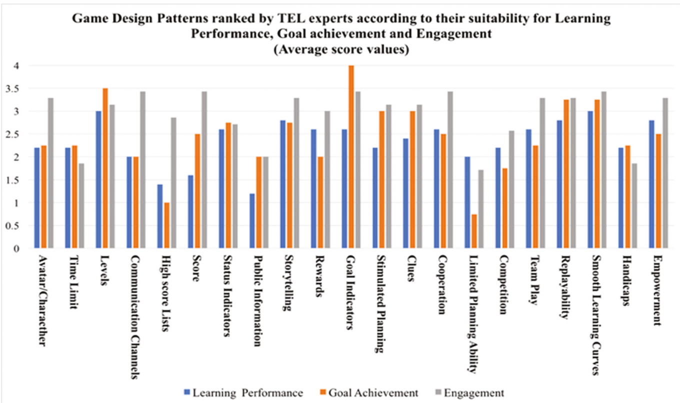


Fig. 3. Representation of GDP average values rated by the TEL Experts related to p1 (blue), p2 (orange) and p3 (grey) (Color figure online)

3.2 Qualitative Analysis

Hints from Game Designers. Each group of game designers was invited to conceptualise the design of a gamified MOOC using, based on their experience, the most suitable game elements to gamify it in order to enhance users' learning performance (p1); goal achievement (p2) or engagement (p3). Each purpose was chosen by two of the six groups, respectively. The topic of the MOOC was given as cyber-security. Each group elaborated a presentation that was performed to the other game designers.

The first two groups that conceptualised the gamification design related to p1 (learning performance) identified the following game elements: Group 1 proposed to use: collaboration via *wiki* and *forum*, aiming at developing a sense of community and information sharing, track of personal progress, *levels* and different levels of *tasks*, with a *rewarding system* for their completion and an inventory for *personal notes*, in which to save helpful posts from the community forum; they also thought of implementing a *game* itself within the MOOC. Group 2 proposed to aim at allowing users to follow an *autonomous path*, as well as a *collaborative path*, that could be enabled by the creation of *alliance*, *asymmetrical information* distribution for the solution of *boss tests*. In addition, group 2 suggested including several *levels* within the MOOC and a *skills tree*[2] which is a game element often present in roleplaying games, (the Diablo[3] series made it famous) and it enables custom configurations of a character's abilities.

Another two groups (3, 4) conceptualised the following game elements for the p2 (goal achievement): Group 3 proposed "*personal profiles* that can be shared with others, *badges* as rewards, *progress bar* and *autonomy*". Group 4 came up with the following idea of transferring the Massive Multiplayer Online Games (MMOG) elements into MOOC, such as: *Skill tree*, connected with the learning needs, learning content and tasks, with a test after each MOOC; "*Knowledge inventory* (completed tasks for the course); *Overview* (whole offer, progress per Skill tree); *Co-op* (Cooperation with "Classes"); *PVP* (Player vs. Player "Knowledge Battle"); *Reward* inside of System (Skill tree, Knowledge Inventory, Succeeded Students as mentor for newbies); *Reward* outside of Systems (Achievements, Link to LinkedIn)".

The last two groups (5, 6) of game designers worked on the conceptualisation of a gamified MOOC for p3 (engagement). Group 5 proposed the use of *competition, collaboration* and *immediate feedback* as game elements to enhance user engagement. They suggested integrating a game concept similar to QuizClash in a MOOC platform: an online quiz game for two players, where one player chooses from several categories and challenges a friend. Both receive the same question and the one who replies faster and correctly wins. The game element was chosen because activating social comparison could engage students. Group 6 proposed using the following game elements to enhance users' engagement in MOOCs: *Quests*, *Narrativ*e, *Player/Character*, *Enemy/Boss*, C*ommunity* (*Guild*)/*Community Experience* and *Status Parameter*. In particular, the narrative conceptualised consists of "some sort of opposing power that

---

[2] It is called a tree because once the basic skills are gained by the users, it opens several branches and the user can decide what to follow.

[3] Blizzard production, 1998. http://eu.blizzard.com/en-gb/games/.

threatens the participants' characters and their private information". In this framework, the participant goal is to work against this power to protect his/her own identity. The players will work together cooperatively against the system and develop a resistance force. As one of the participant states: "The player needs to use what s/he learns in the modules of the course to contribute to the success of this resistance". Being part of this resistance could help in developing "a sense of community similar to MMORPG (Massive Multiplayer- Online Role Play Games) communities such as guilds". "Even if participants are working alone, they should feel that they are contributing to the cause of the resistance/the community", therefore collaboration has to be an option and "it will not be enforced".

Hints from Learning Scientists and TEL Experts. Learning Scientists (LS) as well as TEL experts were not involved in focus groups, however they could express their point of view through the use of the open questions contained in the survey that asked them to detail the advantages and disadvantages of using a specific GDP for the purpose selected. Selected comments are reported here to give a better overview of the LS as well as TEL experts' perceptions on gamification applied to MOOCs.

LS experts gave the following feedback on the three purposes: Starting with the p1 (learning performance), the most rated GDPs were *Levels*, *Empowerment* and with equal score *Avatar/Characters*, *Storytelling* and *Clues*. The advantages listed for these were: *"Levels* give structure in the learning"; *Empowerment* was appreciated because "people like to have autonomy". *Avatar* allows "higher identification with the MOOC", as for *Storytelling*: "Human beings think in the shape of stories, I would guess it clearly enhances performance if done in a way that people can make a meaningful connection with the story" that has been told. Lastly, *Clues* given only at request ("hints button") "could be useful". With respect to p2 (goal achievement), the highest scores recorded were for the following GDPs: *Smooth Learning Curves* and the equally scored *Clues* and *Empowerment*. LS declared that *Smooth Learning Curves* gave the chance to users to track and "experience progress in learning", as for *Clues*, these can work as "scaffolding for learners who need a little more support, through clues everybody can achieve their goals", as disadvantages foreseen: "If it is too easy to obtain clues, the students might not try to figure things out themselves". Regarding *Empowerment*, "it can help users to positively achieve their learning goals". Aiming at p3 (engagement) for LS, the GDPs: *Storytelling*, *Clues* and *Engagement* received the highest scores and the only one commented on was *Clues*: "It helps to have clues, especially for complex goals. However, having them pop up can also distract" the users and be a disadvantage.

TEL experts detailed the following advantages and disadvantages of each GDP for the specific purposes: Starting with p1 (learning performance), the GDPs with the highest ranks were *Levels*, *Smooth Learning Curves* and with equal score *Storytelling*, *Replayability* and *Empowerment*, commented as follows: "*Levels* on the one hand can add motivation to the users (positive) but on the other hand, depending on the users, having several levels to overcome might increase the pressure on the users", "if combined with *time restriction,* this might result in negative effects", "levels can help to monitor progression", "as disadvantages - they would work in long and complex scenarios, which is something to be considered in the MOOC design". With reference to *Smooth Learning Curves*, TEL experts say: "If a learner is an international learner who struggles with language or a novice learner, it may help them through the course"; "apparently only for experts this is not a good strategy, for novices and intermediate learners this should work". About *Replayability*, it can facilitate "learning by failure", anyway it should be given as an option and it "can support the mastering of the learning". As to *Empowerment* advantages: "it can improve engagement" while as a disadvantage it is "often over engineered (serious games), it might lose credibility".

Looking at p2 (goal achievement), the GDPs that recorded the highest scores were *Goal Indicators*, *Levels* and, with an equal score, *Replayability* and *Smooth Learning Curves*. The comments provided by TEL experts on *Goal Indicators* were that it "provides useful insight about a learner's performance and may set the pace of the learning progress, it could especially be useful "as goals might change over time". Concerning *Levels*, an advantage found by TEL experts was: "they can offer flexibility for learners to perform in conditions where they can achieve 'flow' and also set a degree of challenge", however if "the levels of difficulties are badly designed they might affect the learning process" and performance. A disadvantage reported for *Replayability* is that "it might offer learners a way to trick the system into getting higher scores or obtain other rewards", it could however have the advantage of "leading learners to master the concepts". For *Smooth Learning Curves,* the advantage of "avoiding discouragement" among users was reported.

The GDPs *Communication Channels*, *Score*, *Goal Indicators*, *Cooperation* and *Smooth Learning Curves* were selected by TEL experts for p3 (engagement). Comments for *Communication Channels* state that "it is very important to have multiple communication channels to suit different target audience needs. It helps users to feel connected and avoid a sense of isolation"; "Learning is more effective when we engage with peers, so any element to sustain this social component could help if it is used in the right manner". Regarding *Score*: a "personal score system could work in the same way as levels. It helps compete with yourselves rather than others". As for *Goal Indicators* considering that "the success is not defined in MOOCs. One might want to finish only the two weeks that they are interested in. So, if that person puts those goals beforehand, completing them makes that person successful in the course. I think this is very much suitable for the nature of MOOCs". While *Cooperation* on the one hand "can increase engagement", on the other hand "some users will rely on others to reduce effort", it is likely that "it won't help low engaged learners". Lastly, *Smooth Learning Curves* could have as an advantage the decrease of users' "frustration and boredom" but as a disadvantage the TEL experts raise the problem that it is "hard to design".

4 Discussion and Conclusion

Investigating the point of view of game designers, learning scientists (LS) and TEL experts on the selection made on Björk and Holopainen's collection, allows us to understand that despite the different backgrounds of our study participants, there are several points of agreement. With the purpose of identifying suitable GDPs to design our gamification strategy to be applied in a MOOC to enhance users' goal achievement and engagement, we analysed the literature and due to the lamented simplicity of the game elements used up to now and the scarce availability of empirical studies of gamification applied to MOOC, we decided to query other sources, such as those mentioned above [6].

By analysing the data gathered from the questionnaire, it is possible to deduce that among the game designers a common agreement on the evaluation of expected effects of determinate game design patterns within a MOOC for a specific purpose can be recorded. The game designers, for all three purposes, evaluated with a high rank the GDP *Communication Channels*. While *Smooth Learning Curves* was the GDP that received unanimous consensus, but it was selected for only 2 purposes by the game designers, p3 and p1. The LS indicate with a high score the GDP *Clues* for the three purposes; what is more *Empowerment* and S*torytelling* received a high score from the LS but only for the p1 and p2, not for p3. The TEL experts ranked the GDPs: *Smooth Learning Curves* with a high score for all 3 purposes; and *Goal Indicators* for only the purposes p2 and p3.

Considering the similarity between the groups in ranking the GDPs, game designers and LS both chose the GDP *Empowerment* for p2 and related to the p1 and p3 there is not a recorded similarity in the GDPs selected. Game designers and TEL experts issued high ratings related to p1 for the GDP *Smooth Learning Curves*; to p2 purpose for *Goal Indicators* and to p3 for *Communication Channels* and *Smooth Learning Curves*. Comparing LS and TEL experts, on p1 they both rated the GDPs: *Levels*, *Empowerment* and *Storytelling* highly. While for p2 TEL and LS experts ranked the GDP: *Smooth Learning Curves* highly. As for p3 there are no common GDPs with a high score.

With the analysis of the qualitative data we could report and have more clues on the reasons for the scores given as well as having a deeper understanding of the game designers' intention in choosing (or not) a specific pattern for a defined purpose and application. A precious contribution came from the focus group of the game designers with almost a continuous reminder of MMORPG.

Based on this study, our first step towards implementing gamification in MOOC can be concluded by identifying the following GDPs as candidates for further investigation in a field study with the purpose of empirically testing whether the designers' intentions match the effects on users' behaviour stimulating the desired effects or not. Specifically, the GDPs eligible for our implementation are (1) *Empowerment*, (2) *Smooth Learning Curves* (3), *Communication Channels,* (4) *Levels*, (5) *Clues*, (6) *Goal Indicators*, (7) *Skills tree*; (8) *Guild* and (9) *Storytelling*. As the next step, we will design, run and evaluate empirical experiments based on these results.

Acknowledgements. This study is partly funded by the I SECURE - Empowering education systems in information security project (n. 2015-1-IT02-KA201-015005) under the Erasmus+ programme of the European Commission (http://www.isecure-edu.eu/index.php/en/). We would also like to thank the participants who voluntarily took part in this study. A summary of these data can be found in [7].

References

1. Dicheva, D., Dichev, C.: Gamification in education: where are we in 2015? In: E-Learn 2015, Kona, Hawaii, US, pp. 1445–1454 (2015)
2. Dicheva, D., Dichev, C., Agre, G., Angelova, G.: Gamification in education: a systematic mapping study. Educ. Technol. Soc. **18**, 75–88 (2015)
3. The Right to Education | Education | United Nations Educational, Scientific and Cultural Organization. http://www.unesco.org/new/en/right2education
4. Antonaci, A., Klemke, R., Stracke, C.M., Specht, M.: Gamification in MOOCs to enhance users' goal achievement. In: Proceedings of IEEE Global Engineering Education Conference (EDUCON 2017), 25–28 April, Athens Greece. IEEE Xplore (2017)
5. Chang, J.W., Wei, H.Y.: Exploring engaging gamification mechanics in massive online open courses. Educ. Technol. Soc. **19**, 177–203 (2016)
6. Björk, S., Holopainen, J.: Patterns in Game Design (2005)
7. Antonaci, A., Klemke, R., Stracke, C.M., Specht, M.: Identifying game elements suitable for MOOCs. In: Lavoué, É., Drachsler, H., Verbert, K., Broisin, J., Pérez-Sanagustín, M. (eds.) EC-TEL 2017. LNCS, vol. 10474, pp. 355–360. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66610-5_26
8. Lundgren, S., Bjork, S.: Game mechanics: describing computer-augmented games in terms of interaction. In: Proceeding of Technologies for Interactive Digital Storytelling and Entertainment (TIDSE), Darmstadt, Germany (2003)
9. Antonaci, A., Klemke, R., Specht, M.: Towards design patterns for augmented reality serious games. In: Brown, T.H., van der Merwe, Herman J. (eds.) mLearn 2015. CCIS, vol. 560, pp. 273–282. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-25684-9_20

gMOOCs – Flow and Persuasion to Gamify MOOCs

Alessandra Antonaci[1(✉)], Daria Peter[2], Roland Klemke[1,2], Tim Bruysten[2], Christian M. Stracke[1], and Marcus Specht[1]

[1] Welten Institute, Open University of the Netherlands, Heerlen, Netherlands
{Alessandra.Antonaci,Roland.Klemke,Christian.Stracke, Marcus.Specht}@ou.nl

[2] Faculty of Gamedesign, Mediadesign Hochschule, Düsseldorf, Germany
daria.peter93@googlemail.com, T.Bruysten@mediadesign.de

Abstract. Gamification has gained great interest recently in several fields. However, while the literature reports that a gamification design relying on external motivation only can lead users to *cognitive dissonance*, most gamification approaches use points, badges and leaderboards as dominant game elements. We present our developed testable predictions with the aim of investigating additional motivational theories (flow and persuasion) to argue for a deeper integration of gamification and the learning content at hand. Relying on expert selected game elements, we consequently derive design considerations to create gMOOCs, gamified massive online open courses, designed according to the principles of flow and persuasion. Our findings are the basis of our experiment and a contribution to the development of a new theoretical design for gamification.

Keywords: Gamification · Flow · Persuasion · Design · Game elements MOOC

1 Introduction

Gamification in recent years has gained the attention of scholars, researchers and experts from several sectors [1] with its overall aim of helping and fostering behavioural change towards a variety of goals (e.g. better performance, better goal achievement, higher motivation and higher engagement). Gamification is influenced by several disciplines, such as: game design, psychology, instructional design and didactics. This multidisciplinary nature makes its design complex [2]. If gamification is to go beyond merely appealing visualisations, its design implies a deep knowledge and understanding of the context application scenario, the problem/s to be solved, the end users' characteristics (on individual and social connection levels) and of course of game design, in particular the game elements to be transferred.

In our application scenario, Massive Online Open Courses (MOOCs), the problems to be solved comprise a lack of personalisation to reach individual's goals, and a lack of sense of community among MOOC users, both possible causes of high dropout rates,

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 126–136, 2017.
https://doi.org/10.1007/978-3-319-71940-5_12

a phenomenon well-known in these environments. [3]. Today, the game elements most commonly found in implemented gamification solutions are points, badges and leaderboards (known as PBL), which are rewards given to participants who mainly derive their motivational character from social comparison theory [4, 5]. Aiming at generating behaviour change in our target population in order to improve engagement and sense of community, we investigate two further theories: the Flow Theory [6], connected more with engagement at the individual level and the Theory of Persuasion [7, 8], that looks at decision making under social influence in work situations and online communities.

To design a gamification framework, it is fundamental to know and understand how game elements can be inserted in non-gaming contexts. In previous work we thus conducted several studies to select appropriate game elements: we analysed the literature and game element collections such as [9], we involved experts in game design, learning science, and technology-enhanced learning, to assess our first selection of game elements [2, 10] and identified nine game elements rated according to their expected potential to enhance learning performance, goal achievement, and engagement within MOOCs. These elements are: *(g1) empowerment, (g2) smooth learning curves, (g3) communication channels, (g4) levels, (g5) clues, (g6) goal indicators, (g7) skills tree, (g8) guilds, and (g9) storytelling*.

To develop testable predictions, these game elements are better explained and linked to the theories mentioned above to highlight (1) their potential to create or support immersive learning experiences based on flow theory and enhance engagement [11] and (2) their potential to influence learner behaviour based on theories of persuasion [7, 8] and enhance the sense of community (Fig. 1).

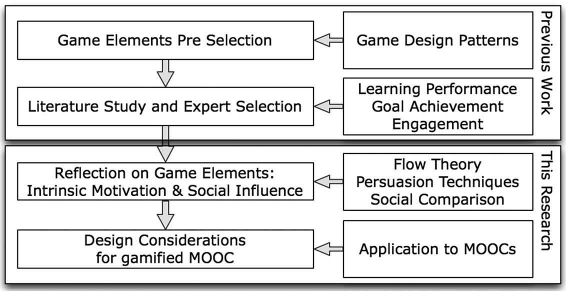


Fig. 1. Research steps towards gamification of MOOCs (gMOOC)

The paper is organised as follows: first the theories of flow and persuasion, are detailed, second a concept of gamification in MOOC (gMOOC) based on these theories is presented. We conclude with our design considerations and future work.

2 Theories of Flow and Persuasion

We refer to these two theories in particular because they allow us to understand how an individual could be engaged in an activity (flow) as well as how the social component can play a role in making choices (persuasion).

According to Csíkszentmihályi, *flow* is "a state of joy, creativity and total involvement, in which problems seem to disappear and there is an exhilarating feeling of transcendence" [12]. This "state" is achievable in types of activities that "have rules that require the learning of skills, they set up goals, they provide feedback, they make control possible. They facilitate concentration and involvement by making the activity as distinct as possible from the so-called "paramount reality" of everyday existence" [12]. Several players can witness to have reached this state.

Csíkszentmihályi also underlines how this pleasant state can be recreated by designing it, flow indeed has a specific set of conditions that have to be met for it to occur. These conditions require a balance of skills and challenges, clear goals, and unambiguous feedback [13] (Fig. 2). In particular, nine components putatively facilitate the flow experience and have to be considered for its design: (f1) *clear goals*; (f2) *focus and concentration*; (f3) *loss of the feeling of self-consciousness*; (f4) *distorted perception of time*; (f5) *immediate feedback*; (f6) *balance between challenge and skill*; (f7) *sense of control*; (f8) *intrinsically rewarding character of activity*; and (f9) *merging of action and aware*ness [14]. Here below they are better described:

(f1) Clear goals. In order for users to get involved in an activity, access to procedural information is needed. Goals facilitate order in consciousness, i.e. they provide a spectrum for action, set expectations, and focus on users' intentions [15].

(f2) Focus and concentration. The execution of a task or skill requires a certain amount of conscious focus [14]. Activities like sports that require a fast shift of attention, and therefore focus, demand a constant evaluation of the situation at hand that is accompanied by periods of low and high focus.

(f3) Loss of feeling of self-consciousness. When a person is in the flow "there is no room for self-scrutiny. Because enjoyable activities have clear goals, stable rules, and challenges well matched to skills, there is little opportunity for the self to be threatened" [11].

(f4) Distorted sense of time. The alteration of time is experienced as either increasing in speed or slowing down [11].

(f5) Immediate feedback. Since flow is a dynamic state of mind, during which a constant improvement of skills takes place, transparency of performance parameters and their adjustment to skills acquired is crucial [11]. If the information on performance is ambiguous or only available when no longer relevant, it may lead to a mental state described as *cognitive dissonance* [16].

(f6) Balance between challenge and skills. This is the most notable dimension of flow that is essential for flow to emerge during an activity. Here, "challenge" is perceived as any opportunity an individual can respond to, and "skill" as the capability of responding to it accordingly, with the ability to actively influence possible outcomes in one's favour [14]. Figure 2 illustrates the 8-channel flow model.

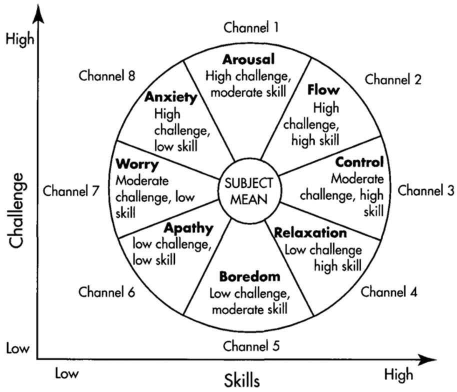


Fig. 2. "A model for the analysis of experience. Perceived challenge is on the ordinate, and perceived skill is on the abscissa" [17].

"The original model assumed that enjoyment would occur along the entire diagonal, that is, when challenges and skills were both very low, as well as when they were both very high. Empirical research findings later led to a modification of the model. People did not enjoy situations in which their skills and the outside challenges were both lower than their accustomed levels. The new model predicts flow only when challenges and skills are relatively in balance, and above the individual's mean level [...]" [11].

(f7) Sense of control. A strengthened sense of control might seem to contradict the loss of self-consciousness, however, it is described as a state of mind during which an individual has no awareness of control but is rather unconcerned about a possible lack of control [18]. Introspectively, the amount of control during flow is perceived as adequate, as an inappropriate amount of control could result in anxiety, thus disrupting flow [14]. In that context, individuals engage in activities that include a certain amount of risk or uncertainty not in the pursuit of thrill or excitement, but because they are competent in influencing the situation at hand.

(f8) Intrinsically rewarding. One of the major characterisations of the flow experience is that its rewards can be derived from the experience itself, making it *autotelic*. It represents a "psychological state, based on concrete feedback, which acts as a reward in that it produces continuing behaviour in the absence of other rewards" [18]. In that sense, intrinsic rewards contain the feelings of personal achievement and joy, being merits that are not primarily associated with material enrichment that strengthen long-term motivation. In contrast, if an activity is mainly providing external rewards like monetary items or objects of value, motivation to endorse oneself further in that particular activity is decreased [19], making voluntary engagements in that activity in the future less likely.

(f9) Merging of action and awareness. In a situation in which all of an individual's relevant skills are needed to comprehend challenges at hand, attention is completely utilised to cope with the activity [11]. Actions then appear to happen spontaneously and automatically, with consciousness immersed in the actions performed.

Continuing our explanation with the second theory: the theory of persuasion (influence) developed by Cialdini, sets out six principles, describing humans taking decisions under social influence [7]. In our life as well as in games and in MOOC environments people are invited to make choices. In investigating online interactions, even if the *non-verbal clues* are not recorded, "social category cues are still available and people may respond to influence appeals based on those cues" [8]. The theory of persuasion explains, based on the following six principles how human behaviour can be (ethically) guided in making decisions and underlines the importance of the social factor in communication [7]: *(p1) commitment and consistency, (p2) reciprocity, (p3) social proof (consensus), (p4) authority, (p5) liking,* and *(p6) scarcity.* While originally intended for direct interpersonal interaction, we aim to analyse and adapt these principles to gamification as detailed below.

(p1) Commitment and consistency. This principle rests on the assumption that humans strive to avoid inconsistency in their behaviour in social situations, which is associated with indecisiveness, confusion and mental instability, whereas consistent behaviour is associated with intelligence, stability and honesty [7]. Consistent behaviour can be turned into automated behaviour, preceding reflective thinking.

In order for a tactic based on commitment to reach its full potential, an individual has to personally commit to a decision or an offer s/he has received [7]. In theory, a personal statement of agreement or a demonstration of compliance alters behaviour associated with the offer or subject at hand in a way that makes it a part of personal identity and therefore, something to defend or sustain.

(p2) Reciprocation. The general idea behind reciprocation is that after receiving help or a gift, whether asked for or not, our associations with it include an undesirable feeling of indebtedness and obligation which demands we give back what we received [7, 20].

(p3) Social proof. One method of coping with a situation at hand is by observing the behaviour of others and adjusting one's own accordingly [7]. Social proof provides information on what to do or on how to act based on a consensus, deducting that some behaviour is appropriate in a certain situation, when the surrounding social environment is displaying the named behaviour [21]. Additionally, the imitation of behaviour is more likely to occur when the model is someone well-perceived, attractive, relatable and familiar.

(p4) Authority. This principle underlines how much people rely on authority, e.g. they "are more likely to give change for a parking meter to a complete stranger if that requester wears a uniform rather than casual clothes" [22]. Applied to MOOC environments, this principle highlights the importance of credibility. Studying the effect of expertise involved in computer-mediated and face to face discussions in online communities, [8] reports that people are more likely to follow the point of view of the authority perceived as expert.

(p5) Liking. People are more willing to comply with a request from someone they like as examined in situations in which, e.g., one's own purchase of an item is based on the suggestion of a familiar, well-perceived person [7]. Additionally, the principle of liking incorporates a concept called the "halo effect" [23], that describes the emergence of a cognitive bias related to a person and objects or people directly associated with him or her, based on previous evaluations of that person.

(p6) Scarcity. The principle of scarcity refers to the idea that “potential loss plays a large role in human decision making” [7]. Opportunities, offers, or items, are pursued more frequently if they display a “potential unavailability”. Consequently, the degree of scarcity of an item can be used as a rule of thumb to estimate its worth, without the need to analyse the item in detail.

3 Gamification Based on Flow and Persuasion

The three dominant game elements used for gamification in education found in the literature are: points, badges and leaderboards (also known as PBL) [24]. A few examples of gamified MOOCs are also available in the literature [2]. From a design perspective using points, badges and leaderboards implies a focus on extrinsic motivation related to social comparison theory [4, 5]. While Cialdini’s principles of persuasion seem to appeal to compliance and identification related to the social dimension, the flow theory allows us to reflect on gamification design from an individual point of view. Furthermore, Csíkszentmihályi’s flow theory represents a closed approach to behaviour alterations via intrinsically rewarding processes. We aim to use both theories in our comprehensive gamification framework reflecting on the potential to facilitate enjoyment and engagement in the learners by addressing extrinsic and intrinsic motivational aspects.

The starting point for this framework is the selection of the nine expert-selected game elements described earlier extended with the list of most commonly used elements (PBL). To reflect these game elements in the light of flow theory and persuasion techniques, we map them with the dimensions of flow and persuasion as described above. This mapping will lead us towards testable predictions on the effects of specific game elements on elements of flow and persuasion. Table 1 summarises the relations between game elements and flow theory/persuasion techniques, which are detailed below.

(g1) Empowerment (“Players feel that they can affect the events and the final outcome of a game” [9]) delivers a *sense of control* (f7) to learners. *Empowerment* also helps to foster the intrinsically rewarding character of actions (f8) by supporting *self-efficacy, as well as* being able to support the *loss of feeling of self-consciousness* (f3) by involving the user in the action itself. *Commitment and consistency* (p1) is supported since *empowerment* leads to individual decisions the learner commits to and to consistent action-response cycles.

(g2) Smooth Learning Curves (enable the chance to “smoothly progress from novice to master” [9]) help to gain *focus and concentration* (f2) as well as a balance between *challenge and skills* (f6), as they avoid overly complex or too simple learning situations. They also support *commitment and consistency* (p1) by designing the learning progress in a consistent way.

(g3) Communication Channels (“the medium and the methods players can use to send messages to other players” [9]), enable clear and *immediate feedback* (f5) by peers and tutors, but also support *reciprocation* (p2) by allowing learners to communicate and exchange favours. *Social proof* (p3) can be supported by *communication channels* as well, since the behaviour of community members can be part of the communication.

Table 1. Mapping game elements with dimensions of flow and persuasion

	(f1) clear goals	(f2) focus and concentration	(f3) loss of the feeling of self-consciousness	(f4) distorted perception of time	(f5) clear and immediate feedback	(f6) balance between challenge and skill	(f7) sense of control	(f8) intrinsically rewarding character of activity	(f9) merging of action and awareness	(p1) commitment and consistency	(p2) reciprocation	(p3) social proof	(p4) authority	(p5) liking	(p6) scarcity
Game Element															
(g1) Empowerment			x				x	x		x					
(g2) Smooth Learning Curves		x				x				x					
(g3) Communication Channels					x						x	x			
(g4) Levels						x				x					x
(g5) Clues	x	x		x	x	x				x			x		x
(g6) Goal Indicators	x							x		x			x		
(g7) Skills Tree						x				x					
(g8) Guild									x				x	x	
(g9) Storytelling		x	x						x					x	
PBL Elements															
(pbl1) Points					x										
(pbl2) Badges	x					x						x	x	x	x
(pbl3) Leaderboards												x			

(g4) Levels ("part of the game in which all players' actions take place until a certain goal has been reached or an end condition has been fulfilled" [9]) can be used to support a balance between *challenge and skill* (f6), since they complement the *smooth learning curve* with a game design approach to design increasing challenges. Levels also support *consistency and commitment* (p1), by representing closed, consistent environments, which the learner commits to finishing. The design of different levels can also provide *scarcity* (p6) by using various level elements in varying frequency.

(g5) Clues ("game elements that give the players information about how the goals of the game can be reached" [9]) are a game element that can be used in a large variety of different ways. Consequently, clues can be used to indicate *clear goals* (f1), to help *focus and concentration* (f2) by directing the learner to relevant aspects, to help dissolve a *distorted perception of time* (f4) by pointing towards breaks or session ends, to give clear and *immediate feedback* (f5), and to support the balance between *challenge and skills* (f6), by guiding a lost learner towards solutions. *Clues* can also underline *consistency* (p1) by directing the learner's attention or represent a*uthority* (p4) by giving direction. *Clues* should be realised using principles of *scarcity* (p6).

(g6) Goal Indicators (give information to players about their current goals in the game), help to indicate, set, and follow *clear goals* (f1). *Goal indicators* also support *commitment* (by fostering goal achievement) and *consistency* (by allowing for long-term oriented goals) (p1). When *goal indicators* are used to guide the learner, they can also represent *authority* (p4). Furthermore, the achievement of a specific goal can be intrinsically rewarding for the learner (f8).

(g7) Skill trees (enable custom configurations of a character's abilities, often organised in branches) support the design of balanced *challenges and skills* (f6) and as *levels* support *consistency and commitment* (p1).

(g8) Guilds ("associations of players who chose to come together to achieve a common goal" [25, 26]) can be used to streamline *clear goals* (f1) and foster *communication and collaboration*, which helps to *support reciprocation* (p2), *social proof* (p3), *linking* (p5), as well as *authority* (p4).

(g9) Storytelling ("the act of telling stories within the game" [9]) can be flexibly used for many contextualising or framing objectives. It offers opportunities to foster *focus and concentration* (p2) by providing an interesting story line. It can support the *loss of feeling of self-consciousness* (f3) by immersing the learner into the story. When storyline and user activities are well integrated, storytelling can support the merging *of action and awareness* (f9).

(pbl1) Points ("numerical representation of the player's success in the game" [9]) can be used to indicate *immediate feedback* (f5).

(pbl2) Badges (the act of telling stories within the game) can be flexibly used to support *clear goals* (f1) or indicate the balance between *challenge and skills* (f6). Badges can also utilise the principle of *scarcity* (p6) and enable *liking* (p5), *authority* (p4), and *social proof* (p3).

(pbl3) Leaderboards ("give players the chance to rank themselves against other players who have previously played the game" [9] or are playing it in parallel) underline the concept of *social proof* (p3).

4 Application to MOOCs – Design Considerations

Since the implementation of a specific game element is a creative process with many choices, we additionally derive design considerations, which should help to apply game elements in a way helpful to support flow and persuasion as desired. In this section, we thus reflect on the application of the previously introduced game elements in the context of MOOCs to find out how far they can help in reaching one or more of the following goals: *(go1) improve community building, (go2) enhance goal achievement and (go3) engagement*.

(g1) Empowerment. The sense of control given to the learner should be represented by the possibility of explicitly stating individual learning goals, which can be followed-up. This way, the learner can continuously assess his/her way towards achieving the stated goal. Empowerment can consequently contribute mainly to (*go2*), with possible side effects on (*go3*).

(g2) Smooth Learning Curves. Keep the learner in continuous learning progress and thus contribute to (go3). In the light of designing a MOOC, smooth learning curves need to be treated already at early stages of the design, since they need to be reflected in the way, the learning content is structured and organised.

(g3) Communication Channels. Engagement is often related to the sense of community *(go1)*. Communication channels are a key way to connect community members and thus contribute to (*go3*).

(g4) Levels help to foster *smooth learning curves* by creating varying difficulty, which consequently contributes to (*go3*). Furthermore, levels can be used to encapsulate specific learning goals in a closed environment: each level should be to some extent self-contained, which contributes to a higher level of individual goal achievement (*go2*).

(g5) Clues can be a design tool to help *smoothen the learning curves* and affect *go2* and *go3*. Indeed, *clues* can be used to reflect on the individual progress towards goal achievement, by indicating e.g. a deviation from the goals originally set (*go2*).

(g6) Goal Indicators are clearly related to (*go2*) and should be used to monitor individually set learning goals and the learner's progress towards achieving them. If set up as a group activity, they can also affect community building (*go1*).

(g7) Skill trees can be a means to help specify individual learning goals (*go2*), e.g. representing required, gained, or targeted skills, probably having an effect on users' engagement (*go3*).

(g8) Guilds directly support (*go1*), by increasing the sense of community and fostering communication, they can impact on (*go3*). However, guilds and their strong social coherence can also be a hindrance to individual goal achievement (*go2*).

(g9) Storytelling can positively impact (*go3*), by creating a motivating context for the learning content. However, storytelling can be problematic in designing self-contained, individual learning units (as required for *go2*), since the ongoing nature of a story might be disturbed.

(pbl1) Points help to represent learning progress and success, and can be seen as external motivators towards (*go2*) and (*go3*).

(pbl2) Badges can be used as individual indicators for progress towards learning goals and thus support (*go2*). If combined with the scarcity criterion, they can foster (*go3*).

(pbl3) Leaderboards allow learners to compare themselves with other learners and can be expected to contribute to (*go3*).

Based on this reflection we can state that the game elements previously selected seem to be supported by both the theories presented in this paper. Further empirical investigations in the context of MOOC are needed.

5 Conclusions

While most of today's gamification approaches rely on game elements related to social comparison, we looked at game elements that are inherently more deeply integrated into the learning progress within MOOCs. Based on flow theory and principles of persuasion, we have reflected on a number of game elements to assess their potential contribution to raise individual motivation (flow) and social interaction (persuasion). We found that with these game elements additional dimensions of motivation beyond social comparison can be supported. Moving towards implementing selected game elements in MOOCs, we have further analysed the possible contribution of these game elements to reaching the described goals (go1–go3) based on implementation intention theory. We found that, while all the different goals can be supported by some of the selected game elements, also conflicts between game elements can arise (e.g. Storytelling contributes to go2, but complicates go3).

As a conclusion, we can state, that more deeply integrating gamification with the learning content at hand can pave the way towards motivating and intriguing learning environments – gMOOCs, which potentially support individual motivation, social interaction, as well as social comparison to address the current shortcomings of MOOCs. Further research is required to empirically assess and measure effects on the goals described as triggered by the different game elements or combinations thereof.

References

1. Hoflinger, P.J., Zimmerling, E.: Monitoring gamification in international patent documents: technology classes, firms and preliminary value indicators. In: 2016 49th Hawaii International Conference on System Sciences (HICSS), pp. 1319–1327. IEEE (2016)
2. Antonaci, A., Klemke, R., Stracke, C.M., Specht, M.: Gamification in MOOCs to enhance users' goal achievement. In: Proceedings of IEEE Global Engineering Education Conference (EDUCON 2017), 25–28 April, Athens, Greece. IEEE Xplore (2017)
3. Cook, S., Bingham, T., Reid, S., Wang, L.: Going massive: learner engagement in a MOOC environment (2015)
4. Christy, K.R., Fox, J.: Leaderboards in a virtual classroom: a test of stereotype threat and social comparison explanations for women's math performance. Comput. Educ. **78**, 66–77 (2014)
5. Mazarakis, A.: Using gamification for technology enhanced learning: the case of feedback mechanisms. Bull. IEEE Tech. Comm. Learn. Technol. **17**, 6–9 (2015)
6. Csikszentmihalyi, M.: Flow: The Classic Work on How to Achieve Happiness. Rider, London (2002)
7. Cialdini, R.B.: Influence: The Psychology of Persuasion. Harper Business, New York (2007)
8. Guadagno, R.E., Cialdini, R.B.: Online persuasion and compliance: social influence on the Internet and beyond. Soc. Psychol. Internet 91–113 (2005)
9. Björk, S., Holopainen, J.: Patterns in game design (2005)
10. Antonaci, A., Klemke, R., Stracke, Christian M., Specht, M.: Identifying game elements suitable for MOOCs. In: Lavoué, É., Drachsler, H., Verbert, K., Broisin, J., Pérez-Sanagustín, M. (eds.) EC-TEL 2017. LNCS, vol. 10474, pp. 355–360. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66610-5_26
11. Csikszentmihalyi, M.: Flow : The Psychology of Optimal Experience. HarperPerennial, New York (1991)
12. Csikszentmihalyi, M.: Flow: The Psychology of Happiness. Rider, London (1992)
13. Csikszentmihalyi, M.: Finding Flow: The Psychology of Engagement with Everyday Life. BasicBooks, New York (1997)
14. Csikszentmihalyi, M., Csikszentmihalyi, I.S.: Optimal Experience: Psychological Studies of Flow in Consciousness (1988)
15. De Kock, F.: The neuropsychological measure (EEG) of flow under conditions of peak performance (2014). http://uir.unisa.ac.za/handle/10500/14359
16. Festinger L.: A theory of cognitive dissonance (1957)
17. Massimini, F., Carli, M.: The systematic assessment of flow in daily experience. In: Csikszentmihalyi, M., Csikszentmihalyi, I.S. (eds.) Optimal Experience: Psychological Studies of Flow in Consciousness, pp. 266–287. Cambridge University Press, New York (1988)

18. Csikszentmihalyi, M.: Beyond Boredom and Anxiety. Jossey-Bass, San Francisco (2000)
19. Deci, E.L.: Effects of externally mediated rewards on intrinsic motivation. J. Pers. Soc. Psychol. **18**, 105–115 (1971)
20. Regan, D.T.: Effects of a favor and liking on compliance. J. Exp. Soc. Psychol. **7**, 627–639 (1971)
21. Pratkanis, A.R.: The Science of Social Influence: Advances and Future Progress. Psychology Press, New York (2007)
22. Influence at Work: Principles of Persuasion. https://www.influenceatwork.com/principles-of-persuasion/
23. Nisbett, R., Wilson, T.: The halo effect: evidence for unconscious alteration of judgments. J. Personal. Soc. **35**, 250–256 (1977)
24. Silpasuwanchai, C., Ma, X., Shigemasu, H., Ren, X.: Developing a comprehensive engagement framework of gamification for reflective learning. In: Proceedings of the 2016 ACM Conference on Designing Interactive Systems - DIS 2016, pp. 459–472. ACM Press, New York (2016)
25. Riegle, R., Matejka, W.: The learning guild: MMORPGs as educational environments. In: Distance Teaching and Learning (2006)
26. Knautz, K., Göretz, J., Wintermeyer, A.: “Gotta Catch ‘Em All” – game design patterns for guild quests in higher education. In: iConference, pp. 690–699 (2014)

Social Engagement in a Digital Role-Playing Game Dedicated to Classroom Management

Guillaume Bonvin(✉) and Eric Sanchez(✉)

CERF, University of Fribourg, P.-A. de Faucigny 2, 1700 Fribourg, Switzerland
{guillaume.bonvin, eric.sanchez}@unifr.ch

Abstract. Classcraft is a role-playing game for classroom management in high schools. Teachers can create teams and assign an avatar to students, as well as points and ‘powers’ as rewards for desired behavior. Classcraft aims to foster players’ social engagement. We conducted a preliminary study on classrooms from Switzerland. The objective aims to characterize the social component of players’ engagement. Our approach is based on the identification of engaged-behaviors. This work is grounded on the idea that players’ engagement encompasses four components (environmental, social, self-component and action). We developed a methodology based on playing analytics to monitor players’ behavior. The detection of socially engaged-behaviors is based on the collection and analysis of players’ digital interactions with kTBS4LA, a playing analytics tools. Different categories of players emerged in terms of social engagement. The data collected shows that social engagement varies across time, classroom or gender. This variation seems linked both to specific game features.

1 Introduction

1.1 Classcraft, a Role Playing Game

Launched in 2014, Classcraft is a digital role playing game dedicated to classroom management [1]. The objective of Classcraft is to transform the classroom into a role-playing game for the duration of the school year. Teachers can create teams and assign an avatar to students, as well as points and ‘powers’ as rewards for desired behavior. In order to acquire powers, the player must demonstrate behavior that is expected of him by the school, such as participating in class, helping other students. The students are warriors, mages or healers and they can buy and use powers that have an impact on real life.

For example, a student who comes five minutes late to class may use the power called “Invisibility”. Consequently, he will not be punished by the teacher. In case if the student does not own this power, the teacher, as game-master, deducts points. A loss of too many points causes death (which means detention in school depending on the rules decided by the game-master). Students can use individual or collaborative powers. For example, mages have the most powerful powers, often benefitting their entire team. The game intends to foster collaboration within students’ teams.

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 137–147, 2017.
https://doi.org/10.1007/978-3-319-71940-5_13

A platform, called Play-Management System [2], allows for the orchestration of the game: creating teams, assigning avatars, displaying the rules or rewarding students. Students are allowed to access to the platform and to personalize their avatar (Fig. 1).

Since Classcraft has been launched, 2.5 million users in 85 countries are registered.

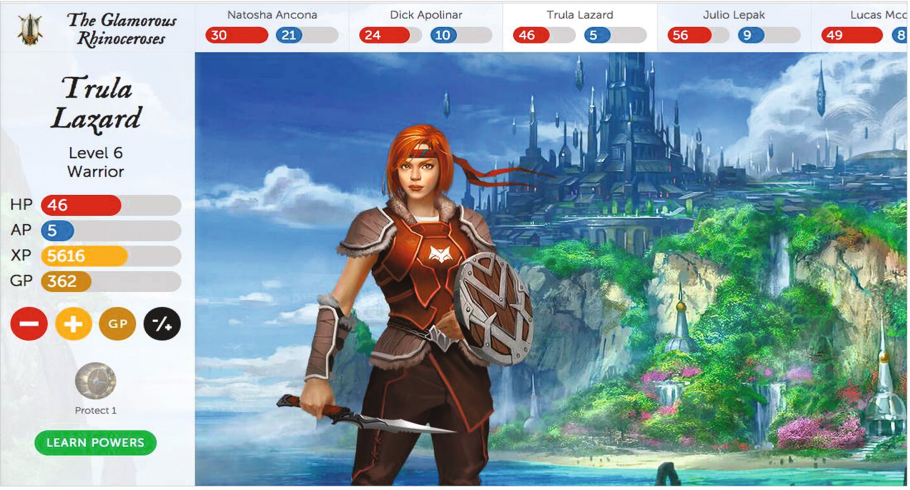


Fig. 1. Classcraft: Student view of avatar, points and level in a team

1.2 Ludicization

Classcraft is a tentative to transform the classroom into a role-playing game. Indeed, different game features are used to convert an ordinary class into a playful situation. These game features have been called by Caillois, *âgon*, *mimicry*, *alea* and *ilinx* [3]. *Mimicry* means playing a role. For classcraft, an avatar with specific characteristics represents each player. *Classcraft* also leverages competition (*âgon*). This competition consists of a conflict with the game itself. Depending on its capacity to adapt his behavior to the rules of the game, the player will win or lose points. Randomness (*alea*) is another gameplay element that is leveraged in Classcraft. Indeed, every class starts with a random event. This event may have a positive or negative impact, such as losing or winning points, on every team. Randomness also manifests itself when, having lost all his points, the player must throw the "cursed die".

However, Classcraft is not limited to use game features in a mechanical way. Classcraft is based on a metaphor that changes the meaning of the actions performed by the students. When Classcraft is played, the classroom is not a classroom anymore. The classroom is a battle where players have to overcome difficulties in order to "survive". An analysis of Classcraft with the Activity Theory framework [4] illustrates that the game enables for a shift of the meaning of the actions performed by the players and not on the action themselves (Table 1).

Table 1. Ordinary class vs classcraft

Levels according to the activity theory	Ordinary class	Classcraft
Operation (How?)	Following the classroom rules (i.e. arriving on time in class)	
Action (Why?)	Driven by the classroom rules (i.e. the teacher expect to have students arriving on time)	Driven by the game rules (i.e. arriving late means losing 10 points)
Activity (For what?)	Being a "good" student (i.e. earning the teacher's esteem)	Being a good player (âgon) (i.e. winning the level)

The shift from an ordinary to a playful situation is not performed at the operation level. However, this shift operated with Classcraft changes the motives of the operation and also the meaning of the activity of the students. Play emerges from the player intention and the meaning given to the performed actions. We use the term *ludicization* [1] to name the changes of motives and meaning of an ordinary class and it conversion into a playful situation.

1.3 Objectives of the Study

By playing Classcraft, students are expected to help other students and to collaborate during school activities. Thus, Classcraft aims to foster players' social engagement. As a result the objective of our study consists of characterizing the players' social engagement and its evolution. We also want to compare social engagement for different classes and to assess if social engagement depends on gender.

2 From Engagement to Social Engagement

Different definitions have been provided for players' engagement [5]. Engagement has been considered to be the lowest level of immersion [6] or a generic indicator of game involvement [7]. Fredricks, Blumenfeld, and Paris [8] consider that engagement encompasses 3 dimensions. *Behavioral engagement* entails participation, positive conduct and effort. *Emotional engagement* relates to the interest of the activity and positive emotions. *Cognitive engagement* entails psychological involvement in the activity and self-regulation. However, due to the ambiguity of the related concepts and their context-dependent definitions [5] the operationalization of players' engagement characterization might be difficult. Thus, our approach is based on the identification of *engaged*-behaviors (*Ibid.*). Based on the Self-Determination Theory [9] which states that motivation results from innate psychological needs, Bouvier et al. [5] consider that players' engagement encompasses four components. The *environmental component* is in relation with the autonomy need, the *self-component* relates to the autonomy need, the *action component* is in relation with the competence and autonomy needs. The *social component* is in relation with relatedness.

3 Methodology

3.1 Playing Analytics

From a methodological point of view, we developed a specific methodology based on *playing analytics* [10] and dedicated to monitor players' behavior. The detection of engaged-behaviors is based on the collection and analysis of players' digital traces. Digital traces are players' interactions performed during a digitally mediated activity. *Obsels* (observed elements) [11] are elementary player's actions (like buying or using powers). Each *obsel* is automatically collected and characterized by a type of event, a timestamp (beginning and end of the event) and information that is useful to derive meaning (attributes and relations with other *obsels*).

Among the wide variety of the collected *obsels*, some are inherent to players' social engagement. As indicators of social engagement we selected two specific categories of *obsels*: buying and using collaborative vs individual powers. Indeed, we consider that the use or the willingness to use collaborative powers brings information on the participation of the player to a collaborative play where the outcomes of the game depend on the capacity of players to take their teammates into consideration.

For this preliminary study, we collected the digital traces produced by secondary students in Switzerland. They played Classcraft during the full duration of the school year.

3.2 Extraction and Data Preparation

The data collected come from 11 classes in Switzerland. For each class, the data consists of 8 JSON files. The events that occurred during the game are described in the main file (logbook.JSON). The following lines give an idea of the format:


```
{
"_id": "qMKyAzyz3zAS6BKMc"
"actionType": "hp",
"actionValue": -10,
"groupID": "BnGfFgigYReGNjtE3",
"targetPlayerID": "bGjjjjF6Jk7nQurh3",
"extraDescription": {
       "description": "matériel",
       "value": -10,
       "_id": "7444699a-9da7-4747-a396-482a07861cdd",
       "mods": {
              "hp": "-10",
        }
},
"screenDisplay": true,

       "userID": "TdTZwQ6ygFKtjBASX",
       "timestamp": "2017-01-18T08:34:08.736Z",
       "createdAt": "2017-01-18T08:34:08.736Z",
```

The event “qMKyAzyz3zAS6BKMc” indicates that the player “bGjjjjF6Jk7nQurh3” (“targetPlayerID”) lost 10 (“actionValue”) HP (“actionType”) because he forgot his school material (“extraDescription”) on January 18, 2017 at 8:30 am (“timestamp”).

All JSON files have been converted to CSV format. *Logbook* file has been completed with information extracted from the other files: gender, team and game-level reached by a given player. These operations were carried out with the digital traces of every class to have the same file structure. Finally, four sets of data (four classes) have been selected for our exploratory study.

3.3 Data Processing with kTBS4LA

The analysis of these digital traces is performed with a specific digital platform called kernel Trace Based System for Learning Analytics (kTBS4LA) [12]. kTBS4LA is a trace-based system platform dedicated to the analysis of digital traces. The data collected are uploaded and different tools dedicated to data processing are available.

First, a kTBS4LA export model is created this model gives a precise description of the digital traces uploaded on the platform. It means that the columns of the file and the different variables are named. The export lasts a few minutes. Once the file is loaded, data exploration can start.

The platform allows for:

- Selecting specific *obsels* from the data collected. The *obsels* that are considered to be relevant for the study (i.e. buying and using collaborative powers) are extracted from the whole dataset.
- The visualization of these *obsels* among a timeline. Different colors and shapes can be used for making apparent specific features (Fig. 2).

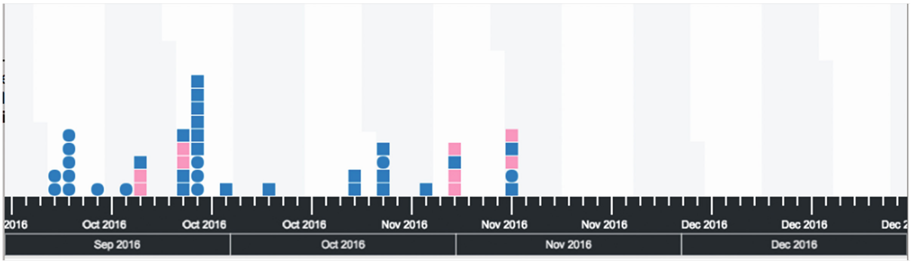


Fig. 2. Example of visualization on the timeline with kTBS4LA. In this view, each obsel is depicted by a square or round dot.

In order to obtain consistent and debatable results, the same protocol is applied to each dataset. We designed and recorded several scenarios for data processing that combine the use of collaborative powers (green) o individual powers (red) with class, genre, teams or avatars. The results of these different scenarios are discussed in the following section.

4 Preliminary Results and Discussion

In this section we discuss the first results of our preliminary study coming from four classes who produced exploitable data. The classes have on average 16 students, 6.75 girls for 9.25 boys and 3.25 teams. The recording of the data last from 4 to 7 months.

Class 4 is a specific case. The teacher gave all the powers to everyone at the beginning of the game. As a result, the players quickly reached a high level (Table 2).

Table 2. Information about the different classes of the study.

Traces	Class 1	Class 2	Class 3	Class 4
Obsels	640 *obsels*	568 *obsels*	2352 *obsels*	7923 *obsels*
Start of data	September 22, 2016 at 16:29	October 13, 2016 at 08:32	15 September 2016 at 16h53	September 22, 2016 at 07:33
Stop of data	December 27, 2016 at 22:30	January 13, 2017 at 09:03	16 January 2017 at 10.10 am	April 09, 2017 at 08:36
Period	3 months and 5 days	3 months	4 months	6 months and 17 days
Average *obsels*/day	17.77	6,1	18, 37	39, 02
Language	French	French	German	German
Students	18 students	13 students	14 students	19 students
Teams	3 teams	2 teams	3 teams	5 teams
Gender	8 girls and 10 boys	3 girls and 10 boys	9 girls and 5 boys	7 girls and 12 boys
Characters	6 Mages 5 Warriors 7 Healers	4 Mages 4 Warriors 5 Healers	3 Mages 6 Warriors 5 Healers	5 Mages 5 Warriors 9 Healers
Use of powers	44 powers	52 powers	184 powers	467 powers

4.1 Evolution of Social Engagement During the School Year

Purchases and uses of powers are two categories of recorded *obsels*. The types of powers purchased or used by players are good indicators of social engagement. Indeed, to progress through the game, students can buy new powers. They can choose between collaborative powers and individual powers. In 3 out of 4 classes, students mostly purchase individual powers (Fig. 3). They first play as individual players.

The data show that, along the school year, the students from classes 1, 2 and 4, use more and more collaborative powers. However, for class 3, this is not the case. There is a balance between individual and collaborative powers. As a result, for 3 classes, social engagement varies along the game. The students are more and more involved into a collaborative play.

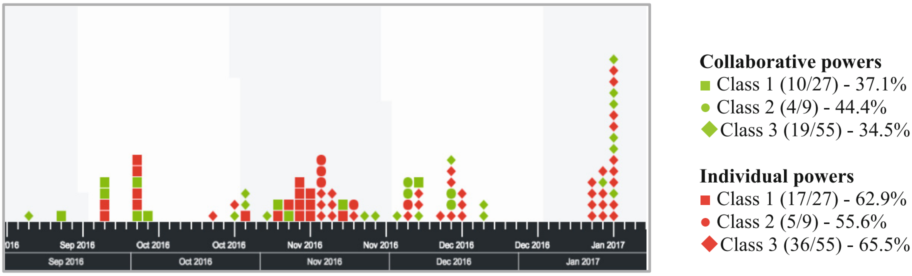


Fig. 3. Purchasing powers: collaborative powers vs individual powers (Color figure online)

The rules of the game intend to foster collaborative behaviors. Indeed, to get through the different game levels, students must use collaborative powers. A majority of students seems to have understood this and, regarding this issue, the game meets the objectives of fostering social engagement in the classroom.

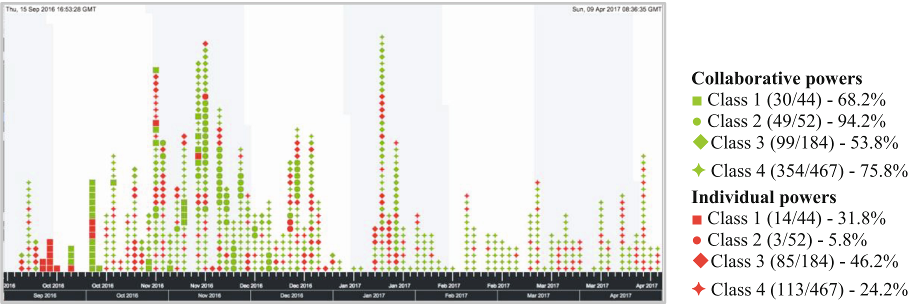
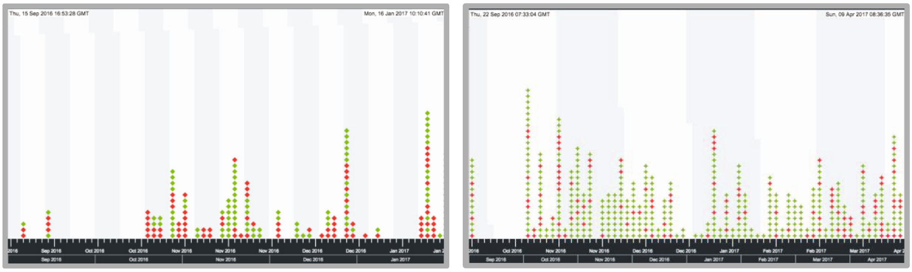


Fig. 4. Using powers: collaborative powers vs individual powers. (Color figure online)

4.2 Comparison of Social Engagement for Different Classes

Figure 5 confirms the evolution of social engagement during the school year. The players use more individual powers at the beginning of the school year. Later on, there is a shift from the use of individual powers to the use of collaborative powers (green dots) (Fig. 4). However, as mentioned above, this is not the case for class 3. The comparison of class 3 and class 4 illustrates the difference between players very active and involved into a collaborative play (class 4) and players less active and mostly involved in an individual play. Social engagement varies among classes and this variation is probably due to the context in which the game is implemented. Indeed, Vandercruysse and Ellen [13] stress that the context is of great importance for the game implementation. In this regard there are probably decisions taken by the teachers that influence students behavior.



Class 3 : ◆(99/184) - 53.8% ◆(85/184) - 46.2%

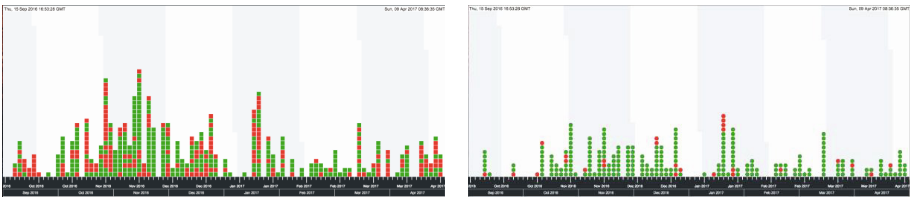
Class 4 : ✦ (354/467) - 75.8% ✦ (113/467) - 24.2%

Fig. 5. Comparison of social engagement for different classes. (Color figure online)

Classcraft changes the classroom context. The class becomes a battle and the game consists of overcoming difficulties in order to survive together. Advancing in the game cannot be done alone. The game itself, but also the rules decided by the teacher as the game-master, converge towards a collaborative approach.

4.3 Gender, Avatars and Social Engagement

The data collected also show differences between how girls and boys play. While boys use more collaborative than individual powers, they still use a lot of individual powers. Girls are less active but more efficient and use few individual power. Figure 6 shows more red squares for the boys than red circles for the girls.



Male
■ collab. (304/493) - 61.7%
■ individual (189/493) - 38.3%

Female
● collab. (228/254) - 89.8%
● individual (26/254) - 10.2%

Fig. 6. Uses by gender: collaborative powers vs individual powers. (Color figure online)

When creating teams, it is recommended that better students (those who are used to respect the rules) play the role of healer or warrior. The healers are supposed to heal their teammates and the warriors to protect the team. Focusing on the data of the 3 teams of class 3, teams 1 and 3 use more collaborative than individual powers, team number 2 uses more individual powers (Fig. 7: Red diamonds).

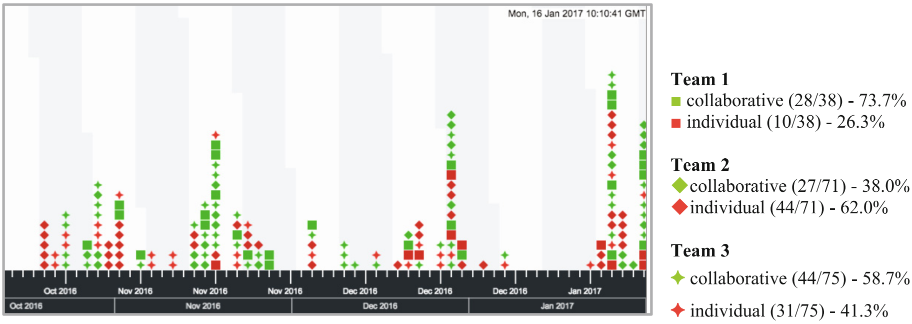


Fig. 7. Uses by team: collaborative powers vs individual powers. (Color figure online)

Table 3 enables for the comparison between mage/warriors and healers. The data collected for team number 2, enables to show the differences on social engagement according to the character played by the students. The 5 players from team number 2 use 71 powers, 27 collaborative and 44 individual. Some students are "strong" healers. They use only collaborative powers. This is the case for two girls. The male students playing as warriors and the female student playing as mage use more individual powers. According to the rules of the games, it means that they are weaker and they need the help of their teammates playing as healers.

Table 3. Collaborative vs individual powers used by players from Team 2 of Class 3

Player 1	**Player 2**	**Player 3**		**Player 4**	**Player 5**
Mage	Warrior	Warrior		Healer	Healer
Girl	Boy	Boy		Girl	Girl
3	5	6		6	7
6	17	21		0	0

Social engagement varies by students' gender. Girls are more involved into collaboration that boys. In addition, social engagement varies among characters played by students.

5 Conclusion

Classcraft aims at fostering desired conduct in students and, based on our preliminary study, there are arguments to state that the game reaches this objective. For three classes that we selected among four from our study, social engagement, assessed through the use of collaborative powers, increases during the time dedicated to play. These first results tend to demonstrate that the game has an influence on how students collaborate with their teammates. The game manages to foster social engagement in students.

It is not totally clear how the game influences students' behavior. Rewards and penalties might play an important role regarding this issue. However, we think that there is no specific game element that can be used to make a game as suggested by previous definitions of gamification [14]. We think that it is possible to subtly combine elements in order to design a context (a metaphor) that changes the meaning of the situation. With Classcraft, the classroom is not anymore an ordinary classroom but becomes a "battle" where students use powers to overcome difficulties. Sanchez et al. [1] use the term *ludicization* to emphasizing that play is performative. They state that it is not possible to "make" the game, as suggested by the suffix "*-fication*" (*facere*) of gamification. As suggested by the suffix "*-icization*, ludicization means that it is possible to convert an ordinary situation into a game. We consider that students' behavior result from the meaning that they give to the experienced situation in the classroom.

The results about students' social engagement seem to support this hypothesis. Students' social engagement varies among gender and role played by the students (warrior, healer or mage). The results also vary among classes. It means that the way the students and the teachers take on the game is of great importance. The differences that we noticed need further investigations and we plan to apply similar data processing frameworks to a larger dataset. We also need to carry out classroom observations and to get a better understanding of the role and influence of the teacher on how students play. We also want to investigate if social engagement developed by the students during the time devoted to play is transferable to ordinary classrooms. In addition to the analysis on a larger dataset, the next step of this study will be a collaborative research carried out with voluntary teachers from a secondary school in Switzerland.

References

1. Sanchez, E., Young, S., Jouneau-Sion, C.: Classcraft: from gamification to ludicization of classroom management. Educ. Inform. Technol. **20**(5), 497–513 (2016)
2. Sanchez, E., Piau-Toffolon, C., Oubahssi, L., Serna, A., Marfisi-Schottman, I., Loup, G., George, S.: Toward a play management system for play-based learning. In: Verbert, K., Sharples, M., Klobučar, T. (eds.) EC-TEL 2016. LNCS, vol. 9891, pp. 484–489. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45153-4_47
3. Caillois, R.: Des jeux et des hommes. Le masque et le vertige. 1967 edn. Gallimard, Paris (1958)
4. Leontiev, A.: Activity, Consciousness, and Personality. Prentice-Hall, Englewood Cliffs (1978)
5. Bouvier, P., Lavoué, E., Sehaba, K., George, S.: Identifying learner's engagement in learning games - a qualitative approach based on learner's traces of interaction. Paper presented at the 5th International Conference on Computer Supported Education (2013)
6. Brown, E., Cairns, P.: A grounded investigation of game immersion. Paper presented at the CHI 2004 Human Factors in Computing Systems, New York (2004)
7. Brockmyer, J.H., Fox, C.M., Curtiss, K.A., McBroom, E., Burkhart, K., Pidruzny, J.: The development of the game engagement questionnaire: a measure of engagement in video game-playing. J. Exp. Soc. Psychol. **45**(4), 624–634 (2009)
8. Fredricks, J., Blumenfeld, P., Paris, A.: School engagement: potential of the concept, state of the evidence. Rev. Educ. Res. **74**, 59–109 (2004)

9. Ryan, R.M., Deci, E.L.: Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. **55**, 68–78 (2000)
10. Sanchez, E., Mandran, N.: Exploring competition and collaboration behaviors in game-based learning with playing analytics. In: Lavoué, É., Drachsler, H., Verbert, K., Broisin, J., Pérez-Sanagustín, M. (eds.) EC-TEL 2017. LNCS, vol. 10474, pp. 467–472. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66610-5_44
11. Cordier, A., Lefevre, M., Champin, P., Mille, A.: Connaissances et raisonnement sur les traces d'interaction. Revue d'Intelligence Artificielle **28**(2–3), 375–396 (2014)
12. Casado, R., Guin, N., Champin, P., Lefevre, M.: kTBS4LA: une plateforme d'analyse de traces fondée sur une modélisation sémantique des traces. Papier présenté lors de l'Atelier ORPHEE Rendez-Vous, Font-Romeu France (2017)
13. Vandercruysse, S., Elen, J.: Towards a game-based learning instructional design model focusing on integration. In: Wouters, P., van Oostendorp, H. (eds.) Instructional Techniques to Facilitate Learning and Motivation of Serious Games. AGL, pp. 17–35. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-39298-1_2
14. Deterding, S., Khaled, R., Nacke, L., Dixon, D.: Gamification: toward a definition. Paper presented at the CHI 2011 Gamification Workshop Proceedings, Vancouver, BC (2011)

OneUp Learning: A Course Gamification Platform

Darina Dicheva(✉), Keith Irwin, and Christo Dichev

Computer Science Department, Winston Salem State University,
Winston Salem, NC 27110, USA
{dichevad, irwink, dichevc}@wssu.edu

Abstract. Gamification of education is still evolving. It lacks systematic studies assessing its effect in different contexts. Creating a gamified course is still time-consuming and design limited. As a response to these challenges, we developed OneUp Learning – a customizable platform aimed at facilitating the process of gamifying learning activities and enabling contextual studies. In this paper we present the platform architecture and its functionality, which includes support for integrating game design elements in learning activities, for creation of dynamic problems and for visualizing student performance and progress. At the end, we present a usability study of the platform and the assessment results from an end user perspective.

1 Introduction

Gamification of learning refers to making learning experiences more engaging and game-like by using game design elements and game mechanics. It increasingly attracts the interest of educators due to its promise to foster motivation and behavioral changes in learning contexts. However, there are two major barriers that hinder the growth of gamified courses: the lack of evidence that gamification will have a positive impact and the lack of appropriate tools to support course gamification. Although the majority of empirical studies on educational gamification have found experimentally more positive than negative effects of gamification on motivation [1–3], its effectiveness for learning remains inconclusive due to limitations of the study design and analysis strategies [4]. The use of gamification in education still lacks systematic and large-scale studies assessing its effect on different learning scenarios with different gamification strategies. A substantial part of the research studies only analyze certain gamification aspects often disconnected from the learning activities. Therefore, the question of how different aspects of gamification actually affect different motivational outcomes has not been empirically addressed and sufficiently understood thus far. All these facts indicate the importance of enabling contextual studies and practices through *customizable educational gamification platforms*.

Although most of the current learning management systems (LMS) offer some gamification support, they are limited both in the assortment of the offered game elements (which are typically not configurable) and in the possible scenarios that can be

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 148–158, 2017.
https://doi.org/10.1007/978-3-319-71940-5_14

gamified. There are also some general gamification platforms, such as Badgeville[1], Bunchball[2], GamEffective[3], Gametize[4], Hoopla[5], and PugPharm[6]. However, the typical approach taken by these platforms is to pack in selected gamification techniques in 'one size fits all' systems, so that they would respond to the needs of enterprises with varying organizational structures. Education and learning, in particular, involves activities different from those in the corporate world, which entails the need of a dedicated platform. This motivated us to develop a holistic gamification platform which enables experimental research on gamifying learning and facilitates the gamification of learning activities with the aim to foster the growth of gamified academic courses.

We are interested in a particular kind of educational gamification – gamification of academic courses or similar types of structured learning experiences. Course gamification is about motivating particular student behavior in the course through the use of game elements, such as instant feedback, freedom to fail, freedom of choice, leveling, progress bars, badges, and leaderboards. Creating a gamified course fitting a specific course structure aligned with the vision of the course designer on the gamified course can be time-consuming and design limited without the support of an adequate platform. Thus, the course gamification platform should be customizable in order to enable instructors to set it up and tailor it to fit their needs. Likewise it should provide support for tailoring gamification features to the preferences of the learners.

Our course gamification platform, named OneUp Learning, is aimed at facilitating the process of gamifying academic courses and enabling tailoring of the gamification features to meet the vision of the course instructor. To the best of our knowledge, it is the first highly configurable platform for gamifying courses or other structured learning activities. In this paper, we describe the platform by presenting the principles, the provided gamification support and the initial results of a usability study.

2 A Platform for Gamifying Academic Courses

The course gamification platform OneUp Learning (OneUp for short) supports the use of established game design principles and elements in the organization of academic courses. Our goal is to support gamification of courses focused on skill development. In such courses students are not merely memorizing the concepts and principles governing the specific discipline, but applying them to everyday life to solve problems using these concepts. Central to the notion of *skill-based learning* is the notion of practice. For developing needed skills in many STEM disciplines learners spend a significant part of their learning time engaged in developing and practicing problem solving and subject-related skills in a variety of hands-on scenarios. Consequently, OneUp enables

---

1 https://badgeville.com/.

2 http://www.bunchball.com/.

3 https://www.gameffective.com/.

4 https://gametize.com/index/.

5 https://www.hoopla.net/.

6 http://www.pugpharm.com/.

instructors to develop exercises for practicing and self-assessment and quizzes or exams for testing particular skills. All learning activities, such as practicing skills, completing quizzes and participating in specified course activities are provided with immediate feedback including detailed progress information and possibly some kind of reward (e.g. points, badges, virtual currency). All elements of the framework - the targeted skills, learning content, activities, quizzes, game design elements and relations between them are configurable, which makes OneUp a course independent and customizable platform. Any game element can be turned on and off to allow studies on the effectiveness of various combinations of elements.

The adopted vocabulary reflects the game terminology: the learning objectives are *skills*, the tests and quizzes are *challenges*, and the questions included in them are *problems*. Earlier ideas about the platform are presented in [5].

2.1 OneUp Functional Model

The OneUp platform's main functionality includes: (1) support for the instructors to integrate game design principles and mechanics in the instructional methods they use in their courses; (2) enabling creation and checking of static and dynamic problems, and (3) learning analytics and visualization to inform students and instructors of the student performance and progress throughout the course. To support all this, the functional model of the platform includes: Course Configuration Tool, Authoring Tool, Gamification Engine, Challenge Tool and Learner Modeling Tool (See Fig. 1).

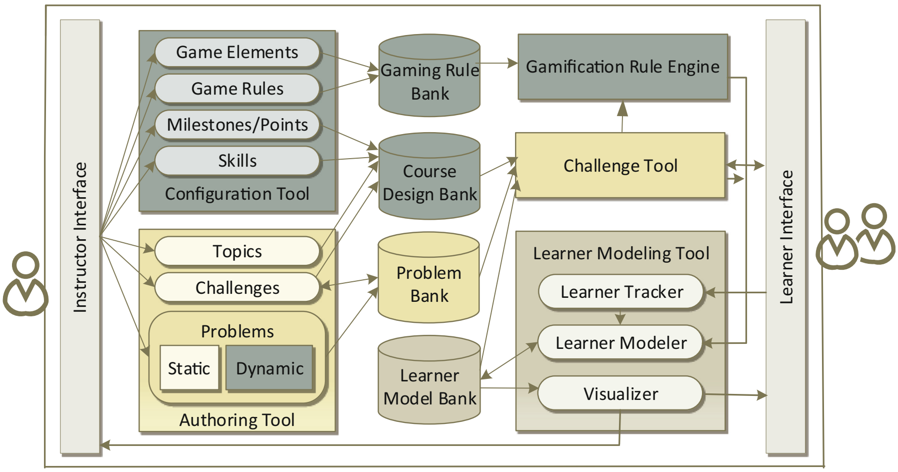


Fig. 1. Functional model of the course gamification platform.

Course Configuration Tool. The OneUp configuration includes two parts: configuration related to the course structure and configuration of the game elements to be used in the particular course. The former includes specifying the course topical structure – the topics to be covered in the course, the learning objectives (skills) targeted in the course, the milestones and activities planned for the course (with their corresponding points), as well adding or importing student data. The second includes the choice of the game elements to be used in the specific course along with the specification of the gaming rules for them. The gaming rules define the conditions upon which certain game elements are applied (e.g. a badge is awarded). The game elements supported by the system are described in Sect. 2.3.

Authoring Tool. The Authoring tool enables entering of problems and challenges. The platform supports two types of challenges: *warm-up challenges* (for student practice and self-assessment) and *serious challenges* (graded course quizzes and tests). When defining a challenge, instructors can choose from problems available in the OneUp's problem bank or create new ones. For each problem included in the challenge, they should specify the *challenge points* earnable from that problem, i.e., the problem's points in the context of the specific challenge (note that the points given for the same problem might be different for different challenges, different courses, or different instructors). The instructor could also specify *skill points* which indicate how the problem contributes to increasing the level of student mastery of related skills (from the set of skills pre-defined for the course). The platform supports two type of problems: static and dynamic, which are described in more detail in Sect. 2.2.

Challenge Tool. The challenge tool delivers the challenges to the students and automatically scores the solutions submitted by them. It returns corresponding feedback to the student and feeds the student results to the learner modeling tool.

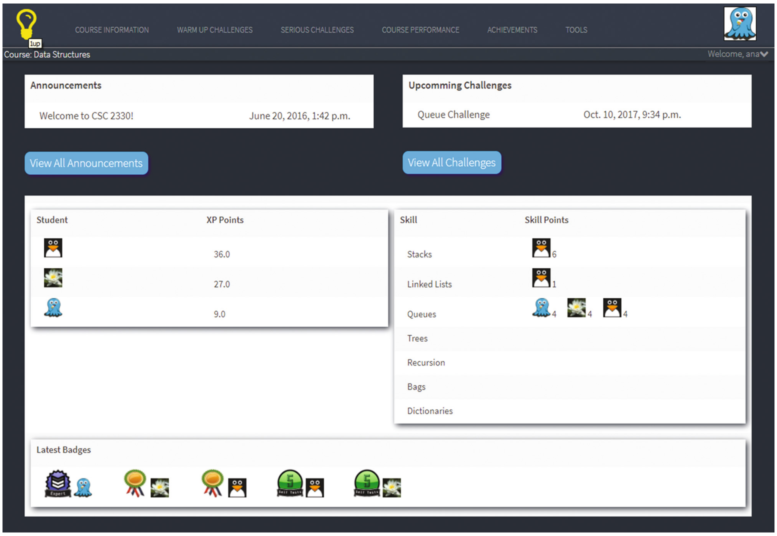


Fig. 2. A screenshot from the OneUp interface.

Learner Modeling Tool. The learner modeling tool consists of three components: learner tracker, learner modeler, and visualizer. The learner tracker collects and processes student transaction logs, providing data for the assessment of student engagement. The learner modeler gets student performance data from the challenge tool, the learner tracker and the game engine and updates the learner model. From the challenge tool it gets the number of attempts, challenge score, correctly/incorrectly answered questions and skill points earned for each course skill associated with this challenge. The game engine informs the learner modeler about awarded badges, earned and spent virtual currency and achieved top status. The learner model is used by the visualizer to display learner performance, progress, and achievements on the student and instructor dashboards as well as on the class leaderboard (see Fig. 2).

2.2 Dynamic Problems

Since freedom to fail and instant feedback are among the fundamental game design principles, in designing OneUp challenges we paid special attention to supporting online practice with automatic assessment. It is important that there will be a sufficient pool of exercises of a particular type for the students to practice. This can be achieved by dynamically generating problem instances from templates, which implies enabling instructors to easily create such templates. OneUp challenges consist of problems which are static and/or dynamic and are scored automatically. *Static problems* are problems for which the correct solution is given at the time of entering them in the system. These include multiple choice questions, multiple answer questions, true/false questions, fill-in-the-blank questions, and matching questions. Problems for which the system does not contain ‘canned’ solutions entered by the instructor we call *dynamic*. These problems are short computer programs which use a random seed to generate a unique instance of a particular programming or calculating problem and then grade the correctness of the answer submitted for that problem.

The platform supports two types of dynamic problems: (1) Programming problems for which students enter a solution in the form of *program code* and the system automatically checks its correctness by running the student’s code through an independent problem solver such as a compiler. For these types of problems, the instructor has to write the problem definition and provide testing code. (2) Problems that are dynamically generated as instances of *parameterized templates*. These templates allow the system to generate multiple non-identical instances of a problem and to check the correctness of the student solutions for them. A template consists of three sections:

- *Set Up Code*: includes Lua code to set up any needed variables and perform any complex math.
- *Problem*: includes the text of the problem. In the text variables are enclosed in [| and |], and Lua code is enclosed in [{and}].
- *Answer*: a call to a function *make_answer* which is used to create answer blanks. Function’s parameters: the name of the answer blank, type, size, answer-checking function (there are several standard answer checking functions provided), and number of points for correct answer in this blank.

2.3 Game Elements Supported by the Platform

We conducted an extensive review of the relevant literature [3, 4] to inform our selection of game elements. As a result we selected the following set:

- *Points*:
  - *Skill points*: These are points indicating the mastery of a specific skill (learning goal) targeted in the course.
  - *Challenge points*: These are total points earned in completing a challenge (e.g. test). The score of a challenge is calculated using the points assigned by the instructor to problems included in the challenge.
  - *Activity points*: These are points given by the instructor for participating in or performing a learning or class-related activity outside of the OneUp system. The instructor is able to define categories of learning activities, such as "attending events", "class presence", "participation in class discussions", etc. when configuring the course and to continuously add activities (with their description and points) to these categories during the semester. The activity categories can be used in game rules and for presenting aggregated information for a student's overall performance in the course.
- *Badges*: The platform offers by default a chain of badges given for mastery of skills: "Novice", "Journeyman", Expert", and "Master". In addition, the instructor is able to define their own badges together with rules for awarding them. For example, "Perfect Score Badge" – for receiving a perfect score on a graded test, "First Challenge Submitter Badge" – for completing a graded challenge first, or "Perfect Monthly Class Attendance Badge". Other examples of conditions include: earning a specified number of points of a particular type; completing a particularly difficult challenge (or passing a specified threshold for a challenge); attempting a certain number of practice exercises; or completing a challenge or learning activity within a certain time frame. A user-friendly graphical interface enables easy entering of a name, description, and image for the badge and specifying rules for awarding it.
- *Levels:* Leveling allows unlocking new problems and challenges based on performance (results from taking challenges, reached skill levels, etc.).
- *Leaderboard*: Supports reputation, social credibility and recognition.
- *Avatars*: Avatars allow learners to assume an identity and express themselves.
- *Virtual currency*: The learners are able to earn course 'bucks' and spend them for course-related benefits following rules defined by the instructor.
- *Immediate feedback*: Apart from the instant reports of the results of taking challenges, the instructor can include additional messages.
- *Freedom to fail*: If a graded challenge is failed, a student may retake it upon certain conditions (defined by the instructor), e.g., after a certain amount of time, after successfully taking a certain number of practice challenges, etc.

The instructors select the game elements that they want to incorporate in the organization of a specific course at the time of creating and configuring it in OneUp.

2.4 Game Engine and Rules

The game engine included in the platform is responsible for executing the rules, which are configured by the instructor and stored in the Gaming Rule Bank. It is event-based and able to fire a chain of rules. Currently we consider the following events: logging into the system, starting a challenge, finishing a challenge, and entering information for a learning activity by the instructor.

The rules specify under what conditions (based on activities, challenges and skills) specific game elements will be applied. For example, a rule can specify that when all three challenges in a certain goal are completed, a certain badge will be given or some content will be unlocked for the student. All rules in the system have the format of a standard production rule:

$$IF < condition > THEN < action > .$$

Here are some examples of rules:

- "*IF a challenge is completed THEN add the appropriate skill points to the skills targeted by the problems in this challenge*."
- *"IF a challenge is completed THEN inform the student how many skill points are needed to reach the next skill badge."*
- *"IF the threshold for a skill badge is reached THEN present the badge to the student."*
- *"IF a challenge is marked as "difficult" and the score is 20% better than the previous challenge taken THEN congratulate the student", etc.*

When certain game elements are selected for a course, some rules will be automatically added for that course. The standard format of the rules allows building an interface for the instructor to define additional rules by using rule templates and selecting choices. Examples of choices are options for using certain rules in the specific course, or numerical values presenting point thresholds, number of attempts, and the like which should be entered in the selected rules' templates. Default rules and values are included to help the instructor in their course configuration.

The OneUp platform can support the gamification of a range of different courses: each course will have its own set of configuration parameters, rules and collection of challenges. Created problems and challenges can be exported and imported in an XML format. This allows their sharing and reuse by different instructors. A novel strategy for reuse which we are developing is the *course gamification design document*. An import of student data from the Blackboard LMS is also supported.

The system is implemented in Python using Django, a Model-View-Controller (MVC) Python Web Framework[7] with a PostgreSQL database. Dynamic problems are implemented in Lua. The interface is in HTML5 and JavaScript, allowing the platform to be used from mobile devices.

[7] https://www.djangoproject.com/.

3 OneUp Usability Study

From the point of view of the end users of an interactive software system, usability is the most important aspect of the system. In terms of gamifying learning activities, Dominguez et al. conclude that a good usability testing process is essential when developing a gamification system, otherwise its motivational effects can be dramatically diminished by unaddressed usability and technical issues [6]. This is even more relevant for general platforms such as OneUp, which are intended for a wide variety of users. If the level of usage of a gamification platform is low because of poor usability, the effect of its engaging elements could be marginalized. In addition, for studying the motivating mechanisms of gamification and measuring the pure effect of gamification, disruptive factors such as poor usability of the gamification platform should be minimized. This section presents the results of usability testing of OneUp that we conducted, using the Software Usability Measurement Inventory (SUMI) [7]. We chose SUMI because it has been extensively validated with a standardization database of software systems with a long track record of success in the commercial market and in university courses [7].

3.1 The Study

Participants. Fifteen participants took part in the anonymous study. (Note that SUMI gives reliable results with as few as 10 users [7].) The subjects were students, instructors and learning technology specialists – a mixture of different types of OneUp end users.

Procedure. The study participants were informed about the purpose of the study and given instructions, consisting of three parts. In the first part, the participants were asked to log into the OneUp platform and complete thirteen typical tasks covering different aspects of the platform usage. Students and instructors were given different tasks, specific to the corresponding user group. Examples of instructors' tasks include creating a serious challenge containing two typical for the course problems (of different type), one of which needs an image import, and configuring the gamification features for the class. Examples of student tasks include taking two warm-up challenges and buying something from the virtual currency course store. In the second part, the participants were asked to go to the SUMI site and complete the SUMI questionnaire. In the third part, they were asked to answer some questions about OneUp aesthetics. The questionnaires were completed anonymously and no additional help was offered.

The SUMI instrument is comprised of 50 questions, in which respondents score each item on a three-point scale (*agree*, *undecided*, *disagree*). Standard scores in SUMI are expressed in a scale where the average is 50, the standard deviation is 10 and higher scores indicate greater satisfaction. So, two thirds of all SUMI standard scores are in the range 40 to 60. Accordingly, the SUMI profiles and analysis of OneUp presented below are computed by using the SUMI mean value 50 and standard deviation 10. A SUMI score is reported for 5 subscales of usability (five different aspects of user satisfaction): *Affect* is the user's general emotional reaction to the software. *Helpfulness* is the degree to which the software is self-explanatory, as well as the adequacy of help facilities. *Control* measures how much the user feels in control of the navigation through the system. *Learnability* measures the facility of mastering the system, that is, how easy it is to learn how to use it. *Global Usability* is a composite measure of usability that describes the user's generalized perceived quality of use.

3.2 Results of the Study

SUMI assesses the five separate aspects of user satisfaction and also includes a global satisfaction scale. Table 1 presents the score of all six scales. It shows that OneUp has high ratings in all five aspects, including the global score. Usability scores for efficiency, affect, control, learnability, and global score are all above or close to 60. Note that a score of 60 is higher than about 85% of the scores in the samples in the SUMI standardized database, thus indicates relatively high user satisfaction. In contrast, a score of 40 is higher than only about 15% of the scores in the normative samples and suggests user dissatisfaction. Learnability received the highest mean score which suggests that for the participants it was relatively easy to pick up how to use the platform. Although the score for helpfulness is lower than remaining scores, it is still above the standardized average.

Table 1. SUMI results of the six scales.

	Mean	St. Dev.	Median	Minimum	Maximum
Global	58.40	10.38	61.0	34	70
Efficiency	60.53	12.01	67.0	37	74
Affect	62.07	12.16	68.0	36	74
Helpfulness	54.67	12.80	56.0	29	70
Controlability	58.73	10.48	60.0	42	74
Learnability	62.20	8.55	65.0	43	71

The above usability results are promising for a prototype system, such as OneUp. The tested interface was liked, and the respondents found it efficient, easy to learn and felt in control. They did not find the platform so helpful, which reflects the fact that the help system had not been completed. The SUMI report contains also a strength and weakness analysis and participant comments on a specific question relevant to the system usability. There were no statistically significant weaknesses of the tested interface in the SUMI report. Answers to the question "What do you think is the best aspect of the platform, and why?" include:

> *"The possibility for the instructor to configure whether they want to use particular game features. The visualization of dashboard elements, badges and achievements is eye catching.", "The leaderboard area is a neat feature. Being able to see how other students are faring in the same class is pretty cool.", "The ability to earn virtual currency is very engaging. Knowing that I can retake an assignment if I had virtual money sounds satisfying."*

The aspects pointing to areas for possible improvements were revealed in some of the answers to the question "What do you think needs most improvement, and why?":

> *"Could add the ability to receive automated hints.", "It would be helpful if there was a quick 'how–to' tutorial section for each activity listed on the home page.", "Make the alternatives (when present) in the form inputs more clear."*

4 Conclusion

The lack of comprehensive understanding of the effects of gamification in learning contexts continues to inhibit educators from adopting effective gamification approaches. While gamification has been regarded as a promising learning technology approach, it has also been acknowledged that it is difficult to design and implement. Furthermore, the effects of gamification have yet to be explored, and the mechanisms behind the positive educational benefits have yet to be isolated and studied. We believe that gamification design must be context specific, where the particular context determines the appropriate combination of game elements for achieving success in gamifying learning. This implies gaining empirical understanding of the relations between particular sets of game elements, types of learners, types of learning activities, and the motivational outcomes. All these challenges entail the need of a platform that can ease the creation of gamified learning activities and support experimental studies in learning context. The OneUp platform was created as a response to that need. It was designed with a twofold purpose: to facilitate the gamification of learning activities and to support experimental studies related to gamifying learning.

The success of any gamified platform is largely dependent on the user motivation and attitude. If a poorly designed interface makes users feel lost, confused, or frustrated, it will hinder its effective use. To evaluate the usability of the OneUp platform we performed a usability study by using the SUMI instrument. The study was intended to provide both an assessment from an end user perspective and also diagnostic information for areas that need improvement.

At the end, we note that successful educational gamification is a collaborative effort of platform designers and instructors. The challenge facing the first group is to create a quality instrument. However, gamifying a learning activity with that instrument so as to be truly motivating and engaging depends on the talent of the instructor and the level of attention given to integrating motivational factors in the learning process.

Acknowledgments. This material is based upon work supported by the NSF project HRD 1623236 "TIP: Increasing Student Motivation and Engagement in STEM Courses through Gamification." We thank Dr. Jurek Kirakowski for being gracious enough to grant us an academic license for SUMI.

References

1. Hamari, J., Koivisto, J., Sarsa, H.: Does gamification work?–a literature review of empirical studies on gamification. In: 47th Hawaii International Conference on System Sciences, pp. 3025–3034 (2014)
2. Seaborn, K., Fels, D.I.: Gamification in theory and action: a survey. Int. J. Hum. Comput. Stud. **74**, 14–31 (2015)
3. Dicheva, D., Dichev, C., Agre, G., Angelova, G.: Gamification in education: a systematic mapping study. J. Educ. Technol. Soc. **18**(3), 75–88 (2015)
4. Dichev, C., Dicheva, D.: Gamifying education: what is known, what is believed and what remains uncertain: a critical review. Int. J. Educ. Technol. High. Educ. **14**, 9 (2017)
5. Dicheva, D., Irwin, K. Dichev, C., Talasila, S.: A course gamification platform supporting student motivation and engagement. In: IEEE International Conference on Web & Open Access to Learning ICWOAL 2014, Dubai, UAE (2014)
6. Dominguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernandez-Sanz, L., Pages, C., Martinez-Herraiz, J.: Gamifying learning experiences: practical implications and outcomes. Comput. Educ. **63**, 380–392 (2013)
7. Kirakowski, J., Corbett, M.: SUMI: the software usability measurement inventory. Br. J. Edu. Technol. **24**(3), 210–212 (1993)

Augmented and Virtual Reality

Design Patterns for Augmented Reality Learning Games

Felix Emmerich[1], Roland Klemke[1,2(✉)], and Thomas Hummes[1]

[1] Faculty of Gamedesign, Mediadesign Hochschule, Düsseldorf, Germany
fnemmerich@gmail.com, Thomas.Hummes@gmail.com
[2] Welten Institute, Open University of the Netherlands, Heerlen, Netherlands
Roland.Klemke@ou.nl

Abstract. Augmented Reality (AR) is expected to receive a major uptake with the recent availability of high quality wearable AR devices such as Microsoft's Hololens. However, the design of interaction with AR applications and games is still a field of experimentation and upcoming innovations in sensor technology provide new ways. With this paper, we aim to provide a step towards the structured use of design patterns for sensor-based AR games, which can also inform general application development in the field of AR.

Keywords: Augmented reality · Game-design patterns
Interaction patterns · Learning games

1 Introduction

"The convergence of mobile computing and wearable computing with augmented reality is naturally of great interest to interaction designers" [1], while "the convergence of wearable computing, wireless networking and mobile AR interfaces" is bringing "a new breed of computing called 'augmented ubiquitous computing'" [2].

Augmented Reality (AR) can be defined as "the fusion of any digital information with physical world settings, i.e. being able to augment one's immediate surroundings with electronic data or information, in a variety of formats including visual/graphic media, text, audio, video and haptic overlays" [3]. Features, "most of which are present in most AR systems" [4] are listed as: "Sense properties about the real world; process in real time; output (overlay) information to the user; provide contextual information; recognize and track real-world objects; be mobile or wearable" [4].

Examples of AR systems utilizing senses other than sight include [5], whose application for cultural sciences students' field trips focused on audio augmentation, arguing that "just like a user should – while driving a car – use sight as much as possible to drive, we believe that with location based learning, a learner's eyes must be primarily used to examine the environment". Haptic feedback corresponding to virtual objects may be transferred from Virtual Reality to AR [6].

Location-based AR outputs information based on the user's position [3, 7, 8]. Points of Interest (POI) are defined and associated with virtual assets – "when a user […] explores a space the POIs are revealed and the content can be accessed" [8]. *Vision-based* AR functions by using computer vision techniques to identify and track

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 161–172, 2017.
https://doi.org/10.1007/978-3-319-71940-5_15

patterns known as *fiducials* (visual markers [9]) in the environment [3, 8]. Both of these approaches have their advantages and disadvantages: Fiducials can only be used with systems trained to recognize them and if conditions like inadequate lighting do not interfere. Location-based systems can suffer from inaccuracy or loss of tracking [10].

A way to combine the advantages of both approaches may lie in hybrid systems as described by [11] or image understanding [12]. The Microsoft HoloLens utilizes a depth camera and tracks head movements through various sensors. A technique called "spatial mapping" [13] is used to construct a three-dimensional model of the surroundings and display virtual content at the appropriate coordinates.

Design patterns for general interaction [14] and games exist [15] and high-level patterns for Mixed Reality games have been proposed [16]. We aim to close the gap between low-level interactions in sensor-based wearable AR systems and high-level game-design patterns for learning games by providing a framework of design patterns for AR games. While such a framework can generally be applied to all kinds of AR games, our main target is to guide the construction of AR learning games. However, pedagogical in relation to the design patterns are not in the focus of this research. Instead, here our focus is more on interactivity and visualization. Definitions, approaches, potentials and limitations of AR are presented, followed by our framework and its first prototypical implementation.

2 Augmented Reality for Learning Games

AR has been applied to many domains, including "hands-free instruction and training, language translation, obstacle avoidance, advertising, gaming, museum tours, and much more" [4], maintenance and repair [17], or Big Data visualization [18], where AR "might solve many issues from narrow visual angle, navigation, scaling, etc.".

Games are an application particularly well-suited for the medium of AR, as "augmented reality is an active, not a passive technology" [7], which emphasizes the "dialogue between the media and the context in which it is used" [3]. Although commercial AR games can be said to go back as far as 2003's EyeToy [10], efforts were for a long time focused on research, until the advance of smartphone technology, which made devices with AR capabilities widely available [16].

Pokémon GO [19], an AR game based on both, the well-known Pokémon franchise and *Ingress* [20], is a rare example of a mobile AR game with a large player base. In the field of learning games [21] reports that Mixed Reality games offer the opportunity to "sense and feel being 'someone' else", while first person experiences make it challenging to develop empathy. *Locatory* is an AR adaptation of the game *Memory*, requiring players to find virtual cards spread around the environment and then match them to real landmarks, to foster orientation skills [22].

Game design patterns have been used to map cognitive and affective learning outcomes in AR games for learning [23]. Similarly, a recent literature review identifies three design principles for learning-oriented AR – "enable and then challenge," "drive by gamified story," and "see the unseen" [24].

Knowledge about how to best approach the design of AR games is still lacking [16], a sentiment [25] shares: "Little is known on how to systematically apply game-design patterns to augmented reality". Similarly to these, [24] attempts to extrapolate design guidelines from the AR game *Dino Dig*, which despite having educational content was primarily intended to entertain.

3 Design Patterns for AR-Based Games

Design patterns describe precisely how to use design techniques in order to achieve certain positive effects, at the same time providing insight and creating a shared vocabulary in the form of a pattern language [16, 26]. More precisely, design patterns "express a relationship between particular design contexts, forces […], and desired ('positive' or good) features" [26].

Björk and Holopainen [27] collected game design patterns, concerned with idea generation. General characteristics of patterns can be outlined as: "Operational and precise"; "positive"; "flexible"; "debatable (the Pattern is clear enough to criticize)"; "testable"; "end-user oriented." [26]. A well-defined game design pattern language would allow for efficient communication, documentation and analysis "e.g. for purposes of comparative criticism, re-engineering, or maintenance" [28].

The patterns collected by Björk and Holopainen [27] do not utilize a problem-solution approach, with Björk, Lundgren, and Holopainen arguing that "not all aspects of design can or should be seen as solving problems, especially in a creative activity such as game design" [15]. The Game Ontology Project describes, analyzes and studies games with pattern-like entries existing in a hierarchy the top level of which includes interface, rules, entity manipulation, and goals [29]. More on the pedagogical side [30] define a framework for the construction of learning games.

The literature revealed only a few pattern approaches for the domain of AR, mainly as *interaction patterns*. The examples below are presented informally but fit characteristics from [26]. They provide data, which the framework presented in this paper was able to expand on and have thus been included.

The *Point Of Interest* (POI) interaction pattern is often implemented in mobile AR browsers. When arriving at pre-defined points, users receive information about the environment through a choice of channels [5]. Browsers may also direct the user towards nearby points of interest [3]. The *Head-Up Display* (HUD) presents information from a fixed point of view, i.e. the information is not assigned some coordinate in 3D space [1]. The *Tricorder* interaction pattern refers to scenarios in which information is scanned from the environment, adding "pieces of information to an existing real-world experience" [1]. *Holochess* experiences consist of presenting entirely virtual objects to the AR environment. *X-Ray Vision*-based experiences allow "seeing beneath the surface of objects, people, or places" [1].

Design patterns for mobile games have been mapped to cognitive and motivational effects in educational AR games [31]. A short, preliminary list of patterns "which take advantage of AR potential" comprising names and short descriptions consists of: *Localization*, *video recording and view sharing*, *synchronous communication*, *contextualization*, and *object recognition* [25]. There are still challenges for AR to overcome, which inform the framework for interactions in AR in the remainder of this paper, applying a pattern approach, incorporating elements from the various sources discussed above, while adhering to the general characteristics laid out by [26].

4 A Pattern-Based Framework for AR Learning Games

Our pattern-based framework, mirroring above-mentioned approaches to game design patterns, is a classification of possible interactions, akin to the game mechanic terminology.

Method. The framework and the software development are based on design patterns, which we defined according to the various approaches presented above [14, 16, 25–28]. Technologically, we build on the comparison of available AR systems and sensors performed by [32]. In this first iteration of the framework, the pattern elements that were used in at least three of the six papers are present. They are:

Name: A succinct name for the pattern.
Forces/Problem: The issues the pattern is intended to combat.
Feature/Solution: A description of one way to solve the problem.
Effects/Consequences: The positive and negative consequences of applying the pattern, including design choices required for implementing the pattern.
Requirements: We introduced requirements, which must or may be met to implement a pattern. This allows game designers interested in implementing patterns to ascertain whether a given pattern fits their criteria.

Scope. Challenges to AR can roughly be sorted into those pertaining to technology, user interface and social acceptance [33]. Due to the scope of this paper and the framework's focus on the interactive medium of games, of these only user interfaces – visualization and interaction – will be covered. Additionally, some patterns focus on the development side of AR applications. The content of the patterns listed below is derived from the literature mentioned, a brainstorming session with participants of the WEKIT project [34], and the characteristics of existing AR games acquired through play testing. We grouped the patterns into six groups: directional, environmental, input, non-visual feedback, media-related, and multi-user, displayed in Table 1.

Table 1. Basic patterns

	Pattern	Forces/Problem	Feature/Solution	Effects/Consequences	Requirements
Directional	Directed gaze	Direct a user's attention to something when they have full control over their view	Use icon to direct user attention. Icon points to the object of interest, if it is not currently visible	Structure AR experience, guide user. Combinable with *Gaze Cursor*. May obstruct other elements and cause screen clutter. Multiple focus points require *Inf. Filtering*	System needs awareness of the focus point position relative to user location and head rotation
	Directed movement	Applications may require the user to move to certain locations	Display icon at the target location. Icon points to the target location, if it is not currently visible	Structures AR experience and guides users. May obstruct other elements and cause screen clutter. Multiple focus points require *Information Filtering*	System needs awareness of the focus point position relative to user location and head rotation
Environmental	Environment adaptation	Design a game that can be run anywhere, while taking the environment into account	Make the game automatically adapt to the characteristics of the environment	The game will run in many environments while making use of AR. Games cannot be planned as stringently and different users may have different experiences	System needs awareness of surrounding. Depth sensors, cameras may be sufficient
	Environment independence	How can you develop AR games for unknown environments?	Make the game not interact with the environment	Games will function more reliably and predictably. User enjoyment may be reduced due to missing AR experience	Games should function without extra sensors
	Environment requirements	Game interacts with different environments without changing the game itself	Make the game's environmental requirements explicit	Ensures that AR game works as intended. Possibly requires user to make changes to their environ-ment to play the game	Requirements must be explained and checked
	Point of interest (POI)	Provide information to users based on location. Bind information to locations not accessible	Bind information to locations, making it available upon getting within a certain range and allowing to direct users to such points	Multiple POIs should not overlap or the system must handle overlap. Avoid screen clutter. Precision is low if only location data is used – may be combined or replaced with a *Gaze POI*	Requires location technology
Input-related	Gaze cursor	How do you select the object to apply actions to?	Base actions on gaze, indicated with visual cursor. Select objects, complete actions (e.g. based on gaze duration)	Guide interactions. Inaccurate or lagging cursors may mislead. Cursor might obscure object of interest. Tunnel vision may be provoked	Awareness of the user's gaze required

(*continued*)

Table 1. (*continued*)

	Pattern	Forces/Problem	Feature/Solution	Effects/Consequences	Requirements
	Gaze point of interest (Gaze POI)	How can events in AR be triggered?	Perform actions when the player looks at specific objects (real or virtual). Events may indicate when something leaves or enters field of view	Focus on important information, context-sensitive information. Combine with *Gaze Cursor* to show selected objects. In larger areas, *POI* may be used. Unintended events may be confusing	Gaze direction and environmental model required
	Gesture based interaction	Interact intuitively with AR environment	Allow the manipulation of objects through gestures	Without special input devices, AR systems may be more immersive and intuitive. Non-intuitive or accidental inputs can be frustrating	Hand movement tracking and gesture recognition required
	Voice commands	How do you allow user input while keeping the user's hands free?	Enables the user to perform actions by speaking appropriate phrases	Can enhance user experience. Simple phrases to avoid frustration. May face acceptance issues in shared spaces	Requires one or more microphones
Non-Visual Feedback	Haptic feed-back	Receive feedback when "touching" augmented objects	Give feedback via the haptic sense. Users may be more responsive to non-visual feedback	Feedback device can limit freedom of movement. Inappropriate feedback may break immersion	Additional feedback channels via feedback devices
Media-Related	Auto-Play	Trigger events in AR without requiring or despite varying kinds of user input	Automatically starting, pausing, and/or stopping events (e.g. based on location/gaze)	Help to avoid tunnel vision and to increase pleasure. Automatic warnings can prevent injury or incorrect use. Can confuse or cause screen clutter. Needs way to stop unwanted content	*Auto-Play* may be based on position, gaze, orientation, audio input, or fiducials
	Information filtering	Screen clutter by too much information. Too little reduces usefulness	Filter information e.g. by distance, angle, or relevance-based information	Makes systems easier to use. Ensure that the filter reflects the intention to avoid too little or too much information	Required sensors depend on the exact filter approach
	Obscured information visualization	Keep track of content that is hidden behind other objects, real or virtual, at a given time	Visualize an object even, or specifically, if it is obscured (various approaches)	Can afford the user a better understanding of their (augmented) environment. However, inappropriate approaches may cause depth perception issues	Tracking of real and AR objects, tracking head and environments

(*continued*)

Table 1. (*continued*)

	Pattern	Forces/Problem	Feature/Solution	Effects/Consequences	Requirements
Multi-User	Shared pointer	How can multiple people in an AR environment communicate efficiently?	Use (gaze) cursors, also visible to other players or allow users to leave markers at set points	Allow users to communicate easily. Requires directed user input (e.g. *Gaze Cursor*). See *Directed Gaze* and *Directed Movement*	Positions and orientations of users must be tracked and synch"d in real-time

5 Combined Patterns for AR Learning Games

The idea of the framework described above is twofold:

(1) We aim to work towards a best practice collection of patterns, which enable players of AR learning games to intuitively understand interaction mechanisms. This aspect has been presented in the review of existing AR games and the patterns found within, which found its place in the effects/consequences column of each pattern.
(2) We want to provide a design toolbox, enabling game designers and developers to construct complex games utilising and combining the basic patterns described. This section is about the second aspect. In line with [27], we see our pattern collection as part of a more general game design language. In this sense, not only should the patterns described here be combinable with each other, but also should a game designer be enabled to combine these patterns with patterns from other collections, such as those reviewed above. Consequently, we see the examples of proposed pattern combinations presented below as a demonstration of the possible use of our framework (Table 2).

Table 2. Combined Patterns usable for learning games

Pattern	Forces/Problem	Feature/Solution	Effects/Consequences
Extended room	The gameplay may require additional virtual rooms, spaces, and structures.	*Obscured Information Visualization*, *Environment Requirements*, and *Environment Adaptation* enable virtual objects to overlay physical structures (e.g. showing non-existing rooms). "Magic Doors" allow to virtually enter a different room.	Useful e.g. for historic learning (what did a scene look like previously) or for change planning (what might it look in future). Also useful for virtual travels.

(*continued*)

Table 2. (*continued*)

Pattern	Forces/Problem	Feature/Solution	Effects/Consequences
Exploration and search	Virtual Objects should be hideable behind/under physical objects to force the player to search them within the physical environment.	Based on *Obscured Information Visualization*, *Information Filtering*, *Environment Requirements*, *Environment Adaptation* from this collection and *Exploration*, *Clues*. Optional hints can be used for guidance.	Explorative learning tasks to foster orientation, exploration strategy
Asymmetric multi-player	Two players require different information, e.g. according to individual roles, locations, or progress.	Based on *Information Filtering*, *Shared Pointer* from this collection and *Asymmetric Information* information can be individualized.	Asymmetric information can be bound to teacher/learner roles or to reflect individual learning needs
Augmented ghost track	A player can see and follow the track of another player in the physical environment	*Shared Pointer*, *Information Filtering*, *Directed Gaze*, *Directed Movement* can be combined to create ghost tracks.	A learner can follow guided steps or work on continuous improvement of own ghost track data.
X-Ray vision	Visualizing internal processes or mechanisms not visible to the eye.	*Obscured Information Visualization*, *Environment Adaptation* and *Gaze POI* are the basis for showing hidden parts	Explain hidden features of complex setups, can be extended to allow for interactivity and animation

We implemented a prototype based on concepts of the open source ARLearn system for mobile serious games [35], adapted for HMDs as a proof of concept. The patterns have been developed for the HoloLens, using the Unity game engine and the HoloToolkit [36]. Some of the patterns (Directed Gaze, Directed Movement, Gaze Cursor, Gesture-based Interaction, and Voice Commands) are already available in HoloLens, others were newly implemented in our prototype (Fig. 1, Table 3).

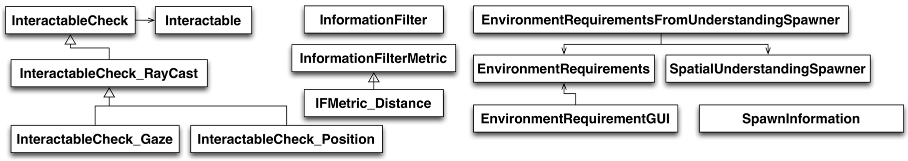


Fig. 1. Pattern framework classes

Table 3. Implemented Patterns

Pattern	Implementation	Comments
Point of interest, gaze POI	*Interactable*, *InteractableCheck_Gaze*, and *InteractableCheck_Position*	The scripts inheriting from *InteractableCheck* provide different ways of detecting *Interactable* objects.
Environment requirements	*EnvironmentRequirements* *EnvironmentRequirementGUI* represents one way of visualizing the available information	Consist of an environmental feature and an upper or lower limit. This is assessed against data from HoloLens.
Environemnt adaptation	Spatial understanding features are utilized in *SpatialUnderstandingSpawner*	Uses predefined sets of rules and constraints from *SpawnInformation* to find spots for instantiating objects, e.g. on a wall/floor, far away.
Information filter	*InformationFilter* executes tasks based on data from *InformationFilterMetric*-derived classes, such as *IFMetric_Distance* (distance between object and player).	The Information Filtering pattern performs actions according to different levels of proximity to the user, but can be extended to cover other metrics.

Interactable, *InteractableCheck_Gaze*, and *InteractableCheck_Position* together form the basis for *Point of Interest* and *Gaze Point of Interest*. The scripts inheriting from *InteractableCheck* provide different ways of detecting *Interactable* objects. The *EnvironmentRequirements* class implements the pattern of the same name. Requirements consist of an environmental feature, an amount, and whether the amount represents an upper or lower limit. This data is compared to that gathered by the HoloLens to assess if the requirements defined on application level are met. *EnvironmentRequirementGUI* represents one way of visualizing the available information.

The spatial understanding features of the device are further utilized in *SpatialUnderstandingSpawner*, an implementation of Environment-Adaption which uses predefined sets of rules and constraints from *SpawnInformation* to find suitable spots for instantiating objects, e.g. on a wall or on a floor, far from the player. *EnvironmentRequirementsFromUnderstandingSpawner* provides a bridge between the two previous mechanisms by generating simple requirements from the parameters of a *SpatialUnderstandingSpawner*.

In first internal usability tests we could show the general functionality and operability of the patterns. In a next step, we aim to integrate the patterns into the general expertise-training framework of the WEKIT architecture [34].

6 Conclusions

We presented an overview of AR definitions, approaches, and applications. We highlighted approaches towards specifying design patterns and created a framework of design patterns for AR games, provided ideas towards the construction of games based on these patterns, and exemplarily adapted a sample of them for Microsoft's Hololens using the Unity game engine.

The work reported here can be expanded in several ways. The framework only covers a fraction of AR interactions, as the scope was limited to user interaction and usability with the Microsoft HoloLens, and it is likely not complete. The results can be seen as a proof of concept. Although focused on learning games, the framework can be used in other AR contexts such as commercial applications, learning applications, or simulations, which aim to make use of AR. Consequently, we see the main contribution of this paper to be a step towards broadening the understanding of AR interaction and application development based on design patterns. As next steps, we aim to apply the framework to the WEKIT AR training solution and to evaluate in pilot application cases related to aircraft maintenace, medical equipment operations, and space craft subsystem integration [34]. Based on these evaluation, we aim to further develop the framework to cover a broad range of AR use cases and interaction scenarios.

Acknowledgments. Parts of this work were supported by the European Commission under the Horizon 2020 Programme under grant agreement No 687669 (WEKIT).

References

1. Lamantia, J.: Inside Out: Interaction Design for Augmented Reality
2. Papagiannakis, G., Singh, G., Magnenat-Thalmann, N.: A survey of mobile and wireless technologies for augmented reality systems (2008)
3. FitzGerald, E., Ferguson, R., Adams, A., Gaved, M., Mor, Y., Thomas, R.: Augmented reality and mobile learning: the state of the art. Int. J. Mob. Blended Learn. **5**, 43–58 (2013)
4. Calo, R., Denning, T., Friedman, B., Kohno, T., Magassa, L., McReynolds, E., Newell, B.C., Roesner, F., Woo, J.: Augmented Reality: A Technology and Policy Primer (2015)
5. Ternier, S., de Vries, F., Börner, D., Specht, M.: Mobile Augmented Reality with Audio. In: Cerone, A., Persico, D., Fernandes, S., Garcia-Perez, A., Katsaros, P., Shaikh, S.A., Stamelos, I. (eds.) SEFM 2012. LNCS, vol. 7991, pp. 53–63. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-54338-8_4
6. Benko, H., Holz, C., Sinclair, M., Ofek, E.: NormalTouch and texturetouch : high-fidelity 3d haptic shape rendering on handheld virtual reality controllers. In: Proceedings of the 29th Annual Symposium on User Interface Software and Technology, pp. 717–728. ACM (2016)
7. Johnson, L., Smith, R., Willis, H., Levine, A., Haywood, K.: The 2011 Horizon Report. The New Media Consortium, Austin (2011)
8. Munnerley, D., Bacon, M., Wilson, A., Steele, J., Hedberg, J., Fitzgerald, R.: Confronting an augmented reality. Res. Lern. Technol. **20**, 39–48 (2012)
9. You, S., Neumann, U.: Fusion of Vision and Gyro Tracking for Robust Augmented Reality Registration (2001)
10. Wetzel, R., McCall, R., Braun, A.-K., Broll, W.: Guidelines for designing augmented reality games. In: Proceedings of the 2008 Conference on Future Play: Research, Play, Share, pp. 173–180 (2008)
11. Schall, G., Wagner, D., Reitmayr, G., Taichmann, E., Wieser, M., Schmalstieg, D., Hofmann-Wellenhof, B.: Global pose estimation using multi-sensor fusion for outdoor augmented reality. In: Proceedings of the 2009 8th IEEE International Symposium on Mixed and Augmented Reality, pp. 153–162 (2009)

12. Furmanski, C., Azuma, R.T., Daily, M.: Augmented-reality visualizations guided by cognition: perceptual heuristics for combining visible and obscured information. In: Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR 2002), IEEE (2002)
13. Spatial mapping. https://developer.microsoft.com/en-us/windows/holographic/spatial_mapping
14. Borchers, J.O.: A pattern approach to interaction design. AI Soc. **15**, 359–376 (2001)
15. Björk, S., Lundgren, S., Holopainen, J.: Game design patterns. In: Level Up: Digital Games Research Conference 2003 (2003)
16. Wetzel, R.: A Case for Design Patterns supporting the Development of Mobile Mixed Reality Games. Found. Digit. Games (2013). http://www.fdg2013.org/program/workshops/papers/DPG2013/b6-wetzel.pdf. Accessed 13 Nov 2017
17. Henderson, S.J., Feiner, S.: Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In: IEEE International Symposium on Mixed and Augmented Reality 2009, pp. 135–144. IEEE (2009)
18. Olshannikova, E., Ometov, A., Koucheryavy, Y., Olsson, T.: Visualizing big data with augmented and virtual reality: challenges and research agenda. J. Big Data **2**, 1–27 (2015)
19. Niantic: Pokémon GO. Game [Android], 6 July 2016. Niantic, San Francisco, CA (2016). Played August 2016
20. Niantic: Ingress. Game [Android], 14 December 2013. Niantic, San Francisco, CA (2013). Played July 2016
21. Kors, M.J.L., Ferri, G., van der Spek, E.D., Ketel, C., Schouten, B.A.M.: A Breathtaking journey. In: Proceedings of 2016 Annual Symposium on Computer Interaction Play - CHI Play 2016, pp. 91–104 (2016)
22. Specht, M., Ternier, S., Greller, W.: Dimensions of mobile augmented reality for learning: a first inventory. J. Res. Cent. Educ. Technol. **7**, 117–127 (2011)
23. Schmitz, B., Klemke, R., Specht, M.: Mobile Gaming Patterns and Their Impact on Learning Outcomes: A Literature Review. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 419–424. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33263-0_37
24. Dunleavy, M.: Design principles for augmented reality learning. TechTrends **58**, 28–34 (2014)
25. Antonaci, A., Klemke, R., Specht, M.: Towards Design Patterns for Augmented Reality Serious Games. In: Brown, Tom H., van der Merwe, Herman J. (eds.) mLearn 2015. CCIS, vol. 560, pp. 273–282. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-25684-9_20
26. McGee, K.: Patterns and computer game design innovation. In: Proceedings of the 4th Australasian Conference on Interactive Entertainment, RMIT University (2007)
27. Björk, S., Holopainen, J.: Patterns in Game Design (2005)
28. Kreimeier, B.: The Case For Game Design Patterns (2002)
29. Zagal, J.P., Mateas, M., Fernández-Vara, C., Hochhalter, B., Lichti, N.: Towards an ontological language for game analysis. In: Proceedings of DiGRA 2005 Conference, pp. 3–14 (2005)
30. Carvalho, M.B., Bellotti, F., Berta, R., De Gloria, A., Sedano, C.I., Hauge, J.B., Hu, J., Rauterberg, M.: An activity theory-based model for serious games analysis and conceptual design. Comput. Educ. **87**, 166–181 (2015)
31. Schmitz, B., Klemke, R., Specht, M.: An analysis of the educational potential of augmented reality games for learning. In: Specht, M., Multisilta, J. (eds.) Proceedings of the 11th International Conference on Mobile and Contextual Learning 2012, pp. 140–147. Helsinki, Finland (2011)

32. Sharma, P., Wild, F., Klemke, R., Helin, K., Azam, T.: D3.1 Requirement analysis and sensor specifications – First version (2016)
33. Azuma, R.T., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B.: Recent advances in augmented reality. IEEE Comput. Graph. Appl. **21**, 34–47 (2001)
34. Klemke, R., Limbu, B., Rasool, J.: WEKIT Framework & Training Methodology – First version (2016)
35. Ternier, S., Klemke, R., Kalz, M., Specht, M.: ARLearn augmented reality meets augmented virtuality. J. Univ. Comput. Sci. Technol. Learn. Across Phys. Virtual Spaces **18**, 2143–2164 (2012). [Special issue]
36. Microsoft: HoloToolkit. https://github.com/Microsoft/HoloToolkit

Learning by Imagining History: Staged Representations in Location-Based Augmented Reality

Peter Winzer(✉), Ulrike Spierling(✉), Erik Massarczyk, and Kathrin Neurohr

Hochschule RheinMain, Unter den Eichen 5, 65195 Wiesbaden, Germany
{peter.winzer,ulrike.spierling,erik.massarczyk, kathrin.neurohr}@hs-rm.de

Abstract. In the SPIRIT project, a location-based Augmented Reality (AR) application has been developed to stimulate the imagination of historical personal life at outdoor places. This paper presents quantitative evaluation results concerning the app's potential for educational purposes. They are based on 107 questionnaires filled by visitors of a Roman fort museum site, having used the AR app in a 40 min tour over several locations. At each place, users can turn into several directions and see keyed video representations of acting characters superimposed on their device's camera image. The fictional events are made up based on historical facts specific to the location. Next to issues with the novelty of the interaction, there has been an interrelation of the subjects self-assessed gaining of knowledge and the appreciation of the story's motivational factors.

Keywords: User experience evaluation · Location-based storytelling Augmented Reality for cultural heritage education

1 Introduction

The application of location-based Augmented Reality (AR) systems in museums and at historical sites has been pursued since the technology has been available. While more often, real objects with attached marker symbols were simply superimposed with information [6, 11], only few examples exist in which live performances are visually blended within the experienced reality [12] without markers. This inevitably leads to novel user interaction patterns, which have been explored in our project 'SPIRIT'. In this paper, we present empirical evaluation results of our playable prototype running on tablets and smartphones, which has been used by regular visitors of an outdoor museum site, the Saalburg Roman fort. The AR app and content was built previously to enable the imagination of the life of typical fort inhabitants during the Roman occupation in Germany. This kind of storytelling of 'staged' events afforded a complex conception of weaving historical facts of this specific location into aspects of family entertainment and dramatic storytelling, as well as addressing still existing limitations of the used AR technology. For technical details, see [7] for the integration of GPS and

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 173–183, 2017.
https://doi.org/10.1007/978-3-319-71940-5_16

markerless image recognition, as well as [10] for our plot engine parsing a formalised content structure (XML) to be authored by game designers, and [16] for the conception of sensor-based interaction patterns, pre-tested with users in about 20 formative evaluation cycles. The resulting prototype has been evaluated in spring 2017 during regular operations of the museum. 107 visitors filled questionnaires after using our tablets in a tour of about 30 to 40 min, being accompanied and observed by two researchers. In the following, we report on selected insights of the quantitative analysis, focusing on potential effects for learning and engagement with history.

2 Related Work

In our project, we created a novel complex interaction style and story format that we have not found elsewhere yet. We evaluated it with regular museum visitors, spanning a great diversity of people including families. The results therefore are unique, as for the most comparable found AR evaluations, subjects had been acquired in a more controlled way [6]. There is still a lack of evaluation in AR with non-technical target groups, as a majority of subjects used to be recruited within the academic field [2] or higher education settings [1].

Besides testing usability, the experience aspired by design relates most to the feeling of presence or sense of place. The latter have been evaluated [4, 14] with applications that so far do not resemble our integration of the search for places, turning around and rendering through video-based storytelling. Still, there are outcomes in line with those of our study, concerning the necessary distribution of attention focus between media and the environment. Recent evaluations addressed learning outcomes and experiential qualities [5, 13]. The results of these support our design goal to not focus on stories for the mere acquisition of knowledge, but for motivational aspects, such as gaining empathy with the past. Additionally, AR does not yet rely on standardized user interaction styles, as different hardware approaches lead to unfamiliar systems that have to be learned by novice users in the first place, which points back to requiring pre-existing motivation of users for success [3, 18].

3 Content and Interactive Experience

One of the first visions of this project, which was inspired by possibilities of location-based Augmented Reality, was to realize the metaphor of ‘meeting the spirits of history’ right where they lived their lives. As the goal was to achieve a feeling of presence of events and action at the historically relevant place, it was found important to tell a story that addresses emotions and imagination, rather than to report facts and numbers. The result of the conception, which included location scouting, historical research and hiring a skilled writer and director, contained several threads of information: (i) A fictional story thread of private love and family life, fitting historical knowledge about the life of inhabitants of the Saalburg fort and associated village, (ii) another fictional story thread imagining events resulting from political relationships between the Roman emperors and the Germanic tribes, including corruption and assaults, and (iii) a list of facts that can be read alongside each fictional scene, backing up the fictional events by approved knowledge, according to the state of history research (Fig. 1, right).

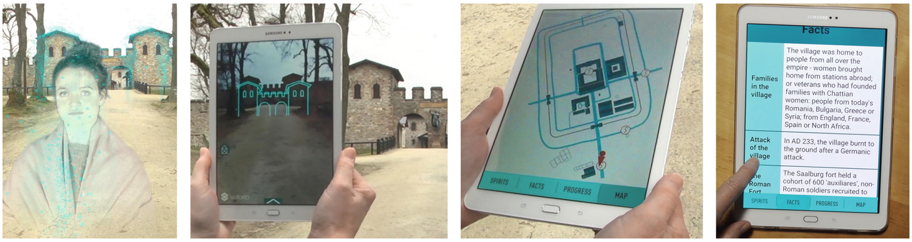


Fig. 1. Left to right: (a) Aurelia indicates the next place by a memory contour. (b) An equivalent memory stencil is used to align the tablet with the main gate. (c) The current GPS position is shown on the map. (d) Brief facts for each scene can be read on demand in a relaxed position.

A special user interaction pattern was conceived, with the goal to let users better experience the on-site reality in which the video representations are blended in. By prompting users to turn around and look into different directions in which scenes are deliberately integrated, the real locality plays an important role as historical setting. As side effect, due to the novel kind of interaction, users have to process more information by learning how to interact successfully. For simplicity and to avoid over-complex interactions in the first instance, the prototype tested worked with a linear story tour.

3.1 Content Description of the Evaluated Experience

The experience was evaluated with regular visitors of the Saalburg [15], a Roman fort partially reconstructed already by the turn of the century AD 1900 with an educational purpose. A brief tutorial explained the interaction with the app. Users follow memories of the spirit "Aurelia", who shows us visual schemes of places that have been meaningful for her. Users can use these as stencils to align the tablet camera to fitting backdrops, for example buildings, found in reality (see Fig. 1, left). Thus, matching background images trigger videos that let appear spirit characters as if seen through a window in reality (Fig. 2). They act out scenes while Aurelia narrates further happenings. Lateral arrows prompt users to turn the tablet 90 degrees to the left or to the right, triggering further 'spirits' in the space around, interacting with each other (see Fig. 2).

Fig. 2. One location with two viewing directions, experienced in a given sequence. (1) Right: To the right of the initial main gate view, two girls appear (in front of low-rise village ruins) with a piece of wood, asking Aurelia to read. (2) Left: Turning the tablet back to the gate after that scene, Aurelia and Aliquander decipher the cry for help from somebody kidnapped.

Aurelia's memories begin in her former village – now only visible as low-rise mural remains – outside the main gate of the Roman fort (Fig. 1, left). After a day of Germanic attacks, two girls – children of a veteran soldier – show her a piece of wood, which expresses a cry for help from a kidnapped cart (Fig. 2). Together with her friend Aliquander, with whom she seems to share neighborly friendship as well as a beginning romance, she follows the carriage. They find out that a Germanic ruler uses corrupt auxiliary soldiers to get hold of Roman weapons, only to beat the Romans with their own means. By finding more spirit-active locations through Aurelia's memory stencils and GPS navigation, we accompany her, sneaking into the fort to warn her father, the Centurion of the fort, in the central building called 'Principia'. When the Romans then try to stop the carriage at the eastern gate, they get into another attack. Here, Aurelia fears for Aliquander's life and realizes that she cannot imagine being without him. Back at the main gate, her memories reveal to the visitor that after this incident, they had married and started a family. With a happy end feeling, they died in their old age, long after the fort had been given up by the Roman Empire as a consequence of the attacks.

Many story events trigger notifications of the "Facts" menu, indicating that factual text information is associated. This can be accessed on demand from the menu, in between fictional scenes. For example, the village was indeed attacked in the year AD 233, and we learn details about the situation of women and private life during the Roman occupation (Fig. 1, right).

3.2 Constraints Influencing the User Experience

Spirits are only visible when the suitable events are triggered – by matching a given GPS position and backdrop image. The software searches all camera input for matching pixel patterns in a set of given reference images. The image recognition (described in [7]) works sufficiently reliable on visually significant buildings. However, on days with frequent weather changes, recognition might fail. Triggering new videos after turns to the left or right is independent of visually significant backgrounds. It is accomplished by the gyroscope sensor, which works incrementally. For novice users, the proper momentum and degree of turning requires tutorial demonstration and some practice. After adopting the movement, it works reliably. The augmented content consists of video sequences with live-acting characters. As expected for AR content, these look like floating in thin air on top of the device's camera image of real surroundings. This is a crucial feature for the experience of presence, but nevertheless, it can lead novice users, who are inexperienced in AR, to miss the visualization when they move the tablet too fast, while the audio stream continues. Concluding, issues due to the technicalities of the prototype and the novel kind of interaction did partially influence the assessed user experience in our test.

3.3 Goals for the Experience and Learning Goals

Although the showcase is in the field of cultural heritage, the acquisition of reproducible factual history knowledge was not the main goal of its design. We wanted to create an imagination of people's personal lives by 'meeting the spirit of history' right at its (former) location. Therefore, the design avoided situations, in which characters also adopt the role of modern museum guides, narrating today's approved knowledge. Thus, explicit history 'learning' is voluntary and takes a back seat to experiencing. The goals we wanted to achieve with the specific AR app design have been (i) the experience of 'presence' of the spirits in the real environment, (ii) motivation and interest to learn more about the Roman world through an emotional story, and (iii) the freedom to access factual information from the menu at the individual user's convenience for appreciating the connection of the story's events to historical knowledge.

4 Empirical Analysis

4.1 Survey

The inquiry was carried out by students of RheinMain University of Applied Sciences, as a supervised learning project, on five days in May/June 2017 at the Saalburg Roman Fort [15]. The interviewers randomly invited regular museum visitors to participate in the testing and survey. 70 groups of visitors could be acquired to walk the tour with the app. Two researchers accompanied each tour, one for support and one observer of situational reactions. 107 subjects from these test groups completed the post-tour questionnaire and answered additional qualitative questions. The questionnaire included 26 groups of questions with 40 single questions, 20 of which were related to the broader subject of 'learning'. Before detailing the statistical analyses, we summarize interesting basic facts (see also Tables 1 and 2):

- Users rated the experience mainly positive (average grades of 4.0–3.8 on a 5-step Likert-Scale, where 5 is highest degree of approval) based on the questions "I had fun using the app", "I wish the app is available also in other museums with fitting content", "I recommend the app".

- 88.8% of the users reported disturbing factors while using the app. These were mainly caused by the queried items of (a) "holding the tablet" (2.8) and (b) "technical problems" (2.7).
- Regarding the questions (a) "I have gained knowledge of the Saalburg when using the app" and (b) "The app motivates me to now learn more about the Saalburg", the users expressed neutral up to rather positive opinions (average grades of 3.1 for (a), and 3.2 for (b)).
- Concerning parameters for the "motivation to continue", all items were ranked rather positive (queried items were "Search for locations" 3.8, "Novelty of the app usage" 3.7, "Suspense of the story" 3.2).

Despite several disturbing factors, the respondents estimate the usage and experience of the application mainly positive. The relationship between the age of the visitors and selected statements about the app is depicted in Tables 1 and 2.

Table 1. Age of visitors and app evaluation (average grades)

Age (years)	Frequency (Age)	App-Experience	Had Fun	Gained Knowledge	Motivated to continue by…		
					Suspense	Novelty	Search
< 20	31.8%	2.1	**4.4**	**3.7**	**3.8**	3.8	**4.3**
20-29	15.0%	3.0	3.3	2.2	2.5	3.3	3.8
30-39	12.1%	2.5	3.5	2.8	2.3	3.8	3.5
40-49	17.8%	1.8	4.2	3.4	3.2	3.4	3.4
50-59	16.8%	1.7	4.1	3.2	3.4	3.7	3.7
60+	6.5%	1.4	3.9	2.3	2.7	**4.0**	3.4
Total	**100.0%**	**2.1**	**4.0**	**3.1**	**3.2**	**3.7**	**3.8**

Younger app users (under 20 years) reported to have the most fun (average grades of 4.4 on a 5-step Likert-Scale) and to have gained the most knowledge (3.7). In contrast, subjects between 20 and 29 years reported to have the least fun (3.3) and to have gained the least knowledge (2.2). This group also reported to have the most experience (3.0 on the Likert scale), concerning the question "I have experience with similar apps (Pokémon Go, Ingress, Geocaching, AR)".

Concerning parameters for the motivation to continue, the "Search for locations" (4.3), and "Suspense of the story" (3.8) were ranked very high by younger people (under 20 years). We conclude that the story is particularly interesting for the younger target group. In contrast, the "Novelty of the App usage" was particularly motivating for senior subjects (4.0). This is likely to be due to the different pre-experiences and expectations of the different age groups. Consequently, particularly for children and teenagers, the application is a new interesting and motivating experience to gain insights and views about the life in the Roman Fort.

For the age group from 40 to 59, mean ratings of between 3.2 and 4.2 are achieved, which makes the app also suitable for this age group. Only for the age groups between 20 and 39, as well as 60+, the app achieved less good ratings, especially in the category "Gained Knowledge".

Table 2. Age of visitors and assessed disturbing factors (average grades, Likert scale)

Age (years)	Frequency (Age)	App-Experience	Disturbing factors while using the app…				
			Noise	Holding Tablet	Other Visitors	Missing Spirits	Technical
< 20	31.8%	2.1	2.0	**1.9**	**1.9**	2.0	2.1
20-29	15.0%	3.0	2.5	3.5	**1.9**	2.8	3.6
30-39	12.1%	2.5	2.2	3.0	2.1	2.6	3.2
40-49	17.8%	1.8	2.6	3.1	2.8	2.5	2.7
50-59	16.8%	1.7	2.5	2.6	1.6	2.1	2.7
60+	6.5%	1.4	2.3	**4.2**	2.0	2.3	2.1
Total	**100.0%**	**2.1**	**2.3**	**2.8**	**2.0**	**2.4**	**2.7**

The subjects were also asked to assess potential disturbing factors by a Likert scale (5 means highest disturbance). Results are listed in Table 2. In general, younger users felt less disturbed by interferences than older users. For the disturbing factors "(Environmental) Noise", "Holding (the) Tablet", "Missing Spirits" (which may be due to technicalities or user errors) and "Technical (Problems)", the average evaluation grades are the lowest in the age group under 20 years. Over 60-year-old users were particularly concerned about holding the tablet as a disturbing factor (average grade 4.2). Here, on the one hand, the lack of habit of this generation in dealing with tablets and, on the other hand, possibly age-related physical limitations are expressed.

Overall, the outlined results show that the evaluated app was particularly suitable for younger people (under 20 years of age), because they (a) had the most fun by using the app, (b) reported to have the greatest learning effect in the app, and (c) felt the least affected by external influences in the app. The elder age groups present more concerns and issues about the usage of the application, however, generally they also vote positively for the application.

4.2 Reliability and Validity

Accordingly to the measurement of reliability with Cronbach's Alpha, in Table 3 only questions related to "Fun/Recommendation", "Learning/Presence/Atmosphere" and "Motivation/Like" are considered as reliable [8]. This implies that only for these 3 topics, the (Cumulative) Variance can be explained by differences in the characteristic to be measured and not by measurement errors, and the results are free of random errors (i.e. the results are reproducible under the same conditions).

Table 3. Validity and Reliability Analysis

Questions related to…	KMO	Bartlett-Test	Cumulative Variance	Cronbach´s Alpha
Fun/Recommendation (a)	**0.652**	$p < 0.000$	**77.637%**	**0.856**
Disturbing Factors/Experience (b)	0.538	$p < 0.000$	**57.717%**	0.533
Learning (c)	0.539	$p < 0.000$	**70.876%**	0.552
Learning/Presence/Atmosphere (d)	**0.811**	$p < 0.000$	**59.429%**	**0.852**
Motivation/Appraisal (e)	**0.694**	$p < 0.000$	**60.093%**	**0.762**

Examples questions for the five topics in Table 3 are: (a) “I enjoyed using the app”, (b) “The following factors bothered me while using the app”, (c) “I have gained knowledge of the Saalburg when using the app” and “The app motivates me to now learn more about the Saalburg”, (d) in addition to (c) “Through the app, I could immerse myself into the Saalburg’s history”, (e) “The following factors motivated me to continue (Story, Novelty, Search)”.

Due to the measurement of the validity with the significant p-values ($p < 0.05$) in the Bartlett-Test and the values bigger than 0.6 in the Kaiser-Meyer-Olkin test (KMO), the concepts “Fun/Recommendation”, “Learning/Presence/Atmosphere” and “Motivation/Appraisal” are valid [8, 9]. Validity refers to the consistency of an empirical measurement with a logical measurement concept. The two concepts (“Disturbing Factors/Experience” and “Learning”), which have not been considered as reliable before, did not create valid results as well. All groups of questions show cumulative variances which are noticeably higher than 50% [9]. This means that between 57.7% and 77.6% of the variances of the collected data can be explained.

4.3 Regression Analysis

For the sake of brevity and the target on learning, we focus in the following on one short regression analysis, which is performed for the dependent variable “I have gained knowledge of the Saalburg when using the app” (Table 4).

Table 4. Regression Analysis – Dependent Variable “I have gained knowledge of the Saalburg when using the app”

Independent Variables/Questions	Regression Coefficient	R-square
Motivation to learn more about Saalburg	0.636	31.6%
Having fun by using the app	0.599	24.4%
Motivation by suspense of the story	0.505	44.1%

The dependent variable “I have gained knowledge of the Saalburg when using the app” is

- significantly positive influenced by the variable “I was motivated by the app to learn more about the Saalburg” (coefficient 0.636, R-square 31.6%),
- significantly positive influenced by the variable “I had fun using the app” (coefficient 0.599, R-square 24.4%),
- significantly positive influenced by the variable “I was motivated to continue by the suspense of the story” (coefficient 0.505, R-square 44.1%).

All the regression coefficients exceed the desired 0.500 (which are classified as sufficient), therefore the values can be classified as sufficient [8, 9]. The R-squares between 24% and 44% show that descriptive variables reach a medium explanation rate.

This means, mainly due to the facts that users (a) have fun with the app and (b) get motivated by the app, they are open to for gaining knowledge about the history of the Saalburg Roman Fort.

4.4 Correlation Analysis

The correlation coefficient analysis determines the degree of the relationship between two individual variables. It is not, however, the degree of dependence, but the degree of the linear relationship/proportionality, which would identify a correlation of 1.000 as 'perfect' relationship. Between the relevant 23 variables/questions, 253 correlation coefficients exist. We identified 65 correlation coefficients which are significant ($p < 0.050$) and have a value over 0.500 (which describes a good relation). For the sake of brevity, we focus on the relationships with correlation coefficients higher than 0.700:

- Relation between variables "I wish the app is available also in other museums with fitting content" and "I recommend the app": (coefficient 0.765), which is almost self-evident.
- Relation between the variables "I was motivated to continue by the suspense of the story" and "I liked the story": (coefficient 0.755), which supports strongly the results of the regression analysis.
- Relation between the variables "I enjoyed using the app" and "All in all, I rate the app concept …": (coefficient 0.718), which shows, that the overall rating of the app is strongly related to the 'fun factor'.
- Relation between the variables "I gained knowledge of the Saalburg through using the app" and "I liked the story": (coefficient 0.708), which shows that the gaining of knowledge is strongly related to the level of appraisal of the story.

5 Conclusion

Within the SPIRIT project, we developed a system and prototypical content for a specialized form of location-based interactive storytelling with Augmented Reality, to support implicit learning by imagining lively events in historical environments. The prototype has been evaluated with end-users concerning some special design goals connected to the AR experience, amongst others, the feeling the presence, as well as the relationship of AR dramatic storytelling with learning.

Based on the quantitative evaluation we found that overall, users liked the app and would recommend it. However, when asked, most users also reported some disturbing factors interfering with their usage, of which some were due to the prototypical nature of the app, but others will most likely remain also in more mature states. In particular, some users (especially seniors) had problems with "Holding the tablet". Younger users felt least affected by external influences. The regression analysis suggests that (a) the parameters "Motivation to learn more about the Saalburg", (b) "Having fun by using the app" and (c) "Motivation by suspense of the story" all have strong influence (regression coefficients: (a) = 0.636, (b) = 0.599, (c) = 0.505) on the extent to which users reported to have gained knowledge of the Saalburg when using the app.

Our learning goal of getting people motivated has yet been achieved mainly for the younger generation. This aspect is under further investigation in the qualitative part of the evaluation.

Acknowledgements. This work has been funded (in part) by the Federal Ministry of Education and Research (BMBF) in Germany (03FH035PA3). We thank all project members for their support (see [17] for videos and personal credits). Special thanks go to students in Media Management at the RheinMain University of Applied Sciences, who conducted the inquiries.

References

1. Bacca, J., Baldiris, S., Fabregat, R., Graf, S.: Kinshuk: augmented reality trends in education: a systematic review of research and applications. Educ. Technol. Soc. **17**(4), 133–149 (2014)
2. Bai, Z., Blackwell, A.F.: Analytic review of usability evaluation in ISMAR. Interact. Comput. **24**(6), 450–460 (2012)
3. Bitter, G., Corral, A.: The pedagogical potential of augmented reality apps. Int. J. Eng. Sci. Invention **3**(10), 13–17 (2014)
4. Chang, Y.-L., Hou, H.-T., Pan, C.-Y., Sung, Y.-T., Chang, K.-E.: Apply an augmented reality in a mobile guidance to increase sense of place for heritage places. Educ. Technol. Soc. **18**(2), 166–178 (2015)
5. Chen, P., Liu, X., Cheng, W., Huang, R.: A review of using augmented reality in education from 2011 to 2016. In: Popescu, E. (ed.) Innovations in Smart Learning. LNET, pp. 13–18. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-2419-1_2
6. Damala, A., Hornecker, E., van der Vaart, M., van Dijk, D., Ruthven, I.: The loupe: tangible augmented reality for learning to look at ancient greek art. Mediterr. Archaeol. Archaeom. **16**(5), 73–85 (2016)
7. Dastageeri, H., Storz, M., Coors, V.: SPIRIT – videobasierte mobile augmented reality lösung zur interaktiven informationsvermittlung. In: Proceedings of DGPF Conference 2015, Band 24, pp. 288–295. DGPF e.V., Köln (2015)
8. Field, A.: Discovering Statistics Using SPSS, 4th edn., pp. 326–345, 684–685, 709. Sage Publications Ltd., London (2013)
9. Hair, J.F.J., Anderson, R.E., Tatham, R.L., Black, W.C.: Multivariate Data Analysis, 3rd edn. Macmillan, New York (1995)
10. Kampa, A., Spierling, U.: Requirements and solutions for location-based augmented reality storytelling in an outdoor museum. In: Culture and Computer Science – Augmented Reality, Proceedings of KUI 2016, pp. 105–117. VWH-Verlag, Glückstadt (2016)
11. Keil, J., Pujol, L., Roussou, M., Engelke, T., Schmitt, M., Bockholt, U., Eleftheratou, S.: A digital look at physical museum exhibits. In: Proceedings Digital Heritage 2013, Marseille, IEEE (2013)
12. Kretschmer, U., Coors, V., Spierling, U., Grasbon, D., Schneider, K., Rojas, I., Malaka, R.: Meeting the spirit of history. In: Proceedings of the International Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST 2001, pp. 161–172. Glyfada, Greece (2001)
13. Li, R., Zhang, B., Sundar, S.Shyam, Duh, H.B.-L.: Interacting with augmented reality: how does location-based AR enhance learning? In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) INTERACT 2013. LNCS, vol. 8118, pp. 616–623. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40480-1_43

14. MacIntyre, B., Bolter, J.D., Gandy, M.: Presence and the aura of meaningful places. Presence Teleoper. Virtual Environ. **6**(2), 197–206 (2004)
15. Saalburg Roman Fort, Archaeological Park. http://www.saalburgmuseum.de/
16. Spierling, U., Kampa, A., Stöbener, K.: Magic equipment: integrating digital narrative and interaction design in an augmented reality quest. In: Proceedings of International Conference on Culture & Computer Science ICCCS 2016, 25–28 October, Windhoek, Namibia (2016)
17. Spirit Homepage. http://spirit.interactive-storytelling.de/
18. Sylaiou, S., Mania, K., Liarokapis, F., White, M., Walczak, K., Wojciechowski, R., Wiza, W., Patias, P.: Evaluation of a cultural heritage augmented reality game. In: Cartographies of Mind, Soul and Knowledge, Special Issue for Prof. em. Myron Myridis, School of Rural and Surveying Engineers, AUTH (2015)

Methods and Tools (for Design and Development)

A Game-Based Development Method of Experiential Learning for Aspiring Professionals

Steven T. de Rooij and Hylke W. van Dijk(✉)

NHL University of Applied Sciences, Leeuwarden, The Netherlands
{steven.de.rooij,h.w.vandijk}@nhl.nl

Abstract. Contemporary workers must demonstrate strong knowledge and skills of their profession. However their so-called 21st century skills such as communication, leadership, and overall teamplay become distinctive in nowadays jobs. Teaching these often tacit skills and attitudes requires an underpinned approach. In this paper we operationalise experiential learning, using a multiplayer game to create the necessary engagement. With this consequential workshop format we created a game-based workshop in which 120 aspiring professionals participated. The results of this exploratory experiment are promising. All participants left the workshop with a clear actionable view to improve team performance. The results of the personality test that is part of the workshop yield an unexpected result. A modified version of workshop may be implemented as a training recommendation instrument for forthcoming students.

Keywords: Agility · 21st century skills · Methodology
Serious gaming · Experiential learning · MBTI

1 Introduction

The nature of work is changing rapidly in our digital society. For one, work in industry and work with service providers becomes more and more data-driven. In this so-called era of smart industry the size of teams is diminishing and routines are in a constant state of flux or even cease to exist [12]. Professionals that work under these circumstances require profound professional knowledge and skills, yet their personal, so-called 21st century skills will become distinctive.

In this paper we focus on personality and organisational context with respect to team performance in changing environments. For an individual to optimally contribute to a team's performance his personality should agree with his supposed role in the team, the job at hand, and the situation in which the job is executed. Individuals require a form of sensitivity and flexibility such that they can adapt their intuitive behaviour into a behaviour that is compatible with a supposed role for optimum team performance. Here we use the Meyers-Briggs personality model [4] as a grounding theory.

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 187–196, 2017.
https://doi.org/10.1007/978-3-319-71940-5_17

Acquiring 21st century skills such as communication, leadership, and overall teamplay, ideally requires one-on-one experimenting and coaching. Although very effective it is not a very efficient way. Modern technology like simulations and games however, allow us to instrument experiential learning where one facilitator works with teams rather than individuals. In this paper we develop a workshop format that operationalises Kolb's experiential learning theory [6,7]. In our approach, learners acquire the necessary personal skills to work in successful teams. This paper reports on the development of the format and on the result of an exploratory experiment with 120 undergraduate students.

While assessing the student's team performance, we examined their individual personality test scores. Much to our surprise the data reveals personality characteristics that differ significantly among groups of students from different training backgrounds. Hence, the workshop format may be applied for extended use, e.g., as a training recommendation instrument for freshman students.

The main contribution of this paper is a structured method for creating workshops, exemplified with a game-based workshop aimed at improving typical 21st century professional skills for teamplay. The method combines contemporary media such as digital questionnaires, simulations, and gaming.

The rest of this paper is structured as follows. Section 2 introduces the underpinning theory, instruments, and related work. Section 3 presents the workshop format development method and experiment. The results of the exploratory experiment are presented in Sect. 4 and henceforth discussed in Sect. 5. Section 6 concludes and comments on future work.

2 Background

Team performance has been studied for many years [13]. Recent publications suggest that personality [2,11] and organisational context [9] are two discriminating predictors for good team performance. Aspiring professionals therefore must be sensitive towards the supposed personality of a role and the situation in which tasks are executed. Additionally it is important to build a form of agility in acting in the scope of the team and its organisational context.

Identifying personality has been a debate for long. Here we choose to use the universally accepted Meyers-Briggs Type Indicator (MBTI) yet the NEO-PI Five-Factor Model [4] could have served our purpose equally well.

The organisational context involves attributes such as hierarchy, roles, and tasks of a team, which obviously depends on the culture and nature of an organisation. It is, however, beneficial for learners to acquire flexibility, so as to adapt to a situation. That is to assess a supposed personality for individual team roles and for individuals to be able to act in a compatible way. Games and simulations offer a proven environment for learners to experience different roles and to assess the efficacy of typical behaviour in a predefined role and contexts [15].

In [14], Thatcher introduces a learning system based on experience and reflection, which builds on Kolb's experiential learning model, see Fig. 1. For learners to appreciate abstract concepts they should interact with the experience of that concept through grasping and transforming. Interactive instruments such as digital questionnaires, simulation, and games provides us with the means to implement such a system. A well defined experience allows learners to grow cognitively as well as affectively such that they value abstract concepts. Note that learners grow individually as well as a team [14].

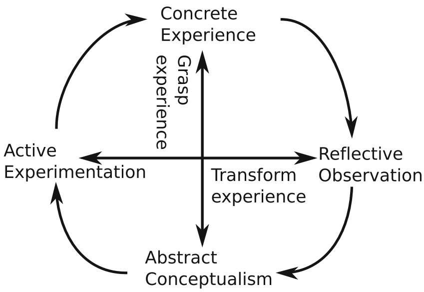


Fig. 1. Kolb's experiential learning model [6].

Both [6,14] emphasise the spiral nature of learning. In fact, Fig. 1 should be regarded as a sequence of entangled circles in which knowledge grows and affection increases with every turn. Following Krathwohl's taxonomy of affective learning stages [8], the first turn creates awareness and willingness to respond to a theory, the second turn provides the learner means to value that theory, and a third turn may yield the learner the ability to organise actions that follow from that theory. Obviously, the instruments need not be identical in every turn. In fact we can exploit the focus on the level of affection and for each turn identify the optimal instrument to interact with the experience.

Simulations and games to further 21st century skills of aspiring professionals have been considered before. Only recently, Coovert et al. [3] report on the use of simulations and games for training and assessing military teams. They structure required characteristics of the learning experience and match that with typical characteristics of gaming. In [5], the authors demonstrate in a video game experiment that team flow is a proper indicator for team performance and the efficacy of learning. For this reason we include an entertainment game in the workshop format. In [10] the authors present a best-practice model for simulation-based measurement of team performance in health care. Their model includes a theory grounded approach, creating multiple perspectives on competences and outcomes, and offering room for debrief, and explicit follow-up.

In order to implement a spiral of experiential learning circles, we rely on a range of instruments. A standard, short version, MBTI test is included to assess personality of subjects and to introduce personality theory in a non-obtrusive manner. We use a quantised interpretation of the MBTI test instead of the usual binary location of a dimension, e.g., in the dimension Sensing/Intuition a subject can score between −4 (Sensing) and +4 (Intuition) whereas ordinarily the score would be either Sensing or Intuition, see also the discussion in Sect. 5.

3 Workshop Development Format

For learners to acquire new skills such as communication, leadership, and overall teamplay we carefully deploy a mix of interactive media and reflection tools. The learning process itself is a spiral implementation of Kolb's experiential learning theory, Fig. 1, where learners gradually increase their level of affection with the theory. Our game-based workshop development format implements Kolb's learning stages cyclically, which are interpreted as follows.

Concrete Experience (CE) yields the observation and representation of facts, i.e., the values shown on a dashboard as a result of the experience.
Reflective Observation (RO) infers the presented facts in the context of the learner, i.e., creating meaning and transfer.
Abstract Conceptualisation (AC) translates the reflection into implications while matching it with theory, i.e., articulating a mission or hypothesis.
Active Experimentation (AE) tests the implications, i.e., gaining the necessary experience to experiment and test a hypothesis.

Substantiating the above workshop development format yields an effective learning workshop. As an exploratory experiment we cast a workshop for aspiring professionals to develop skills such as communication and leadership. The workshop consists of four consecutive rounds, of which we now give a detailed description.

Round 1. Familiarisation with team performance and individuals,
- **CE.** The participants, i.e., team members, feel there is a problem, which they usually state in terms of ill team performance.
- **RO.** Participants share their experience of working in teams and their ability to work on a project.
- **AC.** Participants hypothesise that the personality of individual team members plays a significant role in overall team performance.
- **AE.** In order to identify different personalities within the team, participants fill out a short personality test, e.g., standardised MBTI.

Round 2. Focus on personality of individuals,
- **CE.** Individual test results are presented to all participants.
- **RO.** Inference of test results yields a discussion on similarities and discrepancies of individual test results.
- **AC.** Participants hypothesise that their team performance could be improved by balancing personality, roles, and tasks.
- **AE.** Participants are introduced to the game, i.e., the Artemis Bridge game [1], as a way to reflect on team dynamics by adapting roles and missions within the game.

Round 3. Focus on team performance,
- **CE.** Participants familiarise themselves with roles and tasks that players can fulfil in the game.
- **RO.** Participants define a supposed personality of roles in the context of a chosen mission.

AC. Participants assign roles to team members for the upcoming mission.
AE. Participants play the game in endeavour to accomplish their mission.

Round 4. Focus on transfer,
CE. Participants identify the emerging personalities and behaviours of roles during game play.
RO. Participants assess their ability to adhere to a supposed personality and role.
AC. Focus is returned to team's project and problem. Participants define roles and tasks in their project and define a supposed personality for each role.
AE. Participants return to their project, implementing the experience and insights gained from the workshop.

There is no strict timing associated with rounds and individual stages. It is up to the skilled facilitator to guard time and progress. The above workshop typically is scheduled for a two and a half hour session.

4 Exploratory Experiment

One-hundred-and-twenty ($N = 120$) undergraduate students participated in the exploratory experiment. Participants were invited to sign up for a workshop in teams of five to six persons. Participation was voluntarily, without awarding rewards or credits. Groups signed up to try something new, to develop their team, or simply to engage an interesting experience. As a result, all workshops took place in a natural and relaxed setting with keen students, who were eager to participate. The open nature of the signup method caused that factors like age and gender were not taken into account during evaluation of the results. Participating teams were assigned to one of four groups (Theater, Communication and Multimedia Design (CMD), Technical Business Administration (TBA), and Other) based on their training. The 'Other' group holds teams from a training outside the first three groups.

Round 1. After all participants are seated, the workshop starts with a short introduction in which participants introduce themselves and share their objective for participation. During this time the facilitator introduces topics relevant to the workshop to prepare participants for the first phase. The general role of the facilitator is to observe, to trigger actions, to ask questions, and to provide examples. After the introduction all participants fill in the MBTI personality test on an internet connected tablet.

Round 2. After completion of the personality test, participants present their results to the rest of the group and provide examples of behaviours that support their outcome. The facilitator presents detailed results. For each of the four dimensions of the personality test, each participant scores a value between -4 and $+4$. Where -4 and $+4$ represent a firm preference of such a dimension and zero (0) represents the neutral middle ground.

Participants are given ample time to discuss the results. The facilitator encourages them to provide feedback on the scores of others, in order to complete their view on personality from an external perspective. During the discussion the facilitator gradually shifts focus from individual to team, in order for participants to take a team perspective of their own personality. At the end of this stage the team is aware of different personalities within the team and indicates how these may affect each other.

Round 3. Next, participants are introduced to the Artemis Bridge game and its roles: Captain, Weapons officer, Communications officer, Helmsman, Engineer and Scientist. Participants then determine what personality traits are supposedly present in each of these roles. The goal of this stage is to match individual personalities with supposed personalities of roles in the game. In the process, team roles are assigned for play, participants gain confidence about their choice, and take responsibility for their performance during the game.

Once the Artemis Bridge crew (team) is assembled, the group plays a 10 min introduction round during which the team receives a short explanation of functions and tasks associated with each role. The facilitator encourages participants to familiarise themselves with the game play. After this 10 min introduction play the crew receives a mission instruction for the first round: score as many points as possible. Each enemy destroyed or surrendered increases points, whereas each friendly ship or station destroyed reduces points. Participants get no assistance from the facilitator during this first round.

The first round ends after approximately 15 min. Now the facilitator presents the team with their scores and invites them back to the table to discuss results of the game play. The team is challenged to relate experienced problems during the game to previously identified personalities and behaviours. Such as: “My behaviour in this situation is a good representation of how I would normally behave in the team”. Participants are discouraged to focus too much on positive or negative aspects of their performance. When necessary, the facilitator provides examples of success. The discussion concludes with setting concrete behaviours for the team to increase their performance in a second round of gameplay. In addition participants identify growth by relating behaviours to their personality. Such as: “In order to improve my performance I must alter my behaviour to better fit the supposed personality of my role in the game”.

Round 4. A second 15 min gameplay round ends similarly to the first round, where the facilitator presents the team with the scores of their performance. Participants again discuss their performance but this time the focus is on the results of the objective they set for themselves. That is, reflecting on the effectiveness of the proposed behavioural changes. During this discussion the facilitator gradually shifts focus from the outcomes of the objective in the game to possible outcomes in the real world. The shift allows the participants to transfer experiences and insights gained from the game to the performance of their team and the training project they are engaged in.

4.1 Measurements

During the workshop, apart from the MBTI scores, no other quantitative measurements of growth are recorded. Informally though, teams are without exception excited about the workshop. All teams expressed that they learned valuable lessons from the workshop that they can directly implemented in their current projects. Some of the participants made inquiries whether they could return when they would be working in a new team.

The MBTI scores of the different training groups have been considered more carefully. From our experience during many workshops we learned that participants attribute similar personality traits to certain roles. Participants can fairly accurately describe the distance between their own personality and the supposed personality of a game role. This observation suggest that is possible for participants to assess personal growth in terms of behavioural changes. To pursue this idea further we analysed the MBTI scores of all participants. We expect students from different training background to demonstrate different personalities.

The sample ($M_{\text{age}} = 27.29$, $SD = 11.46$) consists of 84 men and 36 women. A one-way MANOVA was used in order to determine if there were any significant differences between the groups Theater, $n = 12$, Technical Business Administration (TBA), $n = 29$, Communication and Multimedia Design (CMD), $n = 50$ and Other, $n = 28$, for their scores on any of the four MBTI dimensions. The results show that there is a significant difference for the dimensions Extraversion/Introversion $F(3, 115) = 2.73$, $p = .05$ and Sensing/Intuition $F(3, 115) = 3.25$, $p = .02$. The results show no significant differences for the dimensions Thinking/Feeling ($p = .35$) and Judging/Perceiving ($p = .07$). (p = .35) and Judging/Perceiving (p = .07). Further analysis of the multiple comparison tables (Tables 1, 2, 3 and 4) show that the Theater group scored significantly higher on the Sensing/Intuition dimension than any other group; Table 2. Furthermore, for the Judging/Perceiving dimension, the Theater group also scored significantly higher than the TBA group ($p = .05$), but not different than the other two groups; Table 4. While comparisons between the other groups showed

Table 1. Multiple comparison table for dimension Extraversion/Introversion.

Group	cf.	M	p
Theater	TBA	−0.62	0.86
	CMD	−1.14	0.42
	Other	−2.00	0.06
TBA	CMD	−0.52	0.77
	Other	−1.38	0.12
CMD	Other	−0.86	0.40

Table 2. Multiple comparison table for dimension Sensing/Intuition.

Group	cf.	M	p
Theater	TBA	2.57	0.02
	CMD	2.25	0.03
	Other	2.33	0.04
TBA	CMD	−0.32	0.94
	Other	−0.24	0.98
CMD	Other	−0.08	0.99

Table 3. Multiple comparison table for dimension Thinking/Feeling.

Group	cf.	M	p
Theater	TBA	1.16	0.27
	CMD	0.90	0.44
	Other	0.88	0.53
TBA	CMD	−0.26	0.93
	Other	−0.28	0.94
CMD	Other	−0.02	1

Table 4. Multiple comparison table for dimension Judging/Perceiving.

Group	cf.	M	p
Theater	TBA	2.03	0.05
	CMD	1.16	0.39
	Other	1.12	0.41
TBA	CMD	−0.87	0.35
	Other	−0.82	0.52
CMD	Other	−0.05	1

no significant differences, strong indications, however, are found for differences in the other two dimensions (Extraversion/Introversion and Thinking/Feeling); Tables 1 and 3.

5 Discussion

Underlying Kolb's theory is the awareness of growth and the affordance to create a path from a current situation to a desired situation. To facilitate awareness, a workshop should assist a participant in accurately determining his starting point, acquiring knowledge about the underlying processes and theory, and defining his desired endpoint. In the Artemis Bridge game workshop the starting point is explicitly determined through the score on the MBTI test, whereas the endpoint is implicitly set through the supposed personality of a role in the game. The difference between their own position and that of the supposed role can be used as a guideline for growth. In order to effectively gain awareness of these positions a tool is necessary to measure these otherwise subjective starting and endpoint.

The short version of the regular MBTI always presents 1 out of the 16 available profiles. Which profile is assigned is determined by the position that participants have on each of the four dimensions: Extraversion/Introversion, Sensing/Intuition, Thinking/Feeling, and Judging/Perceiving. Although this leads to a wide variety of profiles, detailed information is left out, which could be used to more accurately describe a participants personality. For example, a sales representative and a manager should share traits of the extraversion dimension. However one might argue that one or the other would benefit from more extraversion. Or in terms of our workshop, when a participant shares the outcome of the MBTI with a supposed role, a more accurate position within the dimension would allow for indicating differences. We achieved this by replacing the regular profile outcome with a numeric value between -4 and $+4$ for each dimension.

Our modified interpretation of MBTI scores provides a more precise definition of otherwise subjective outcomes, i.e., defining exact positions of the start and endpoint. For example if the participants outcome on the modified MBTI would be a 0 on the Extraversion/Introversion dimension, and the supposed position would be -2 we can apply this information, while in the original version of the MBTI the outcomes would have been identical. While this modification has proven a useful tool in our workshops, we realise that more research is required to determine its validity.

The modified MBTI test leads to a more concrete personality characterisation. Yet, additional steps are required in order for participant to be able to implement these results of the workshop. One suggestion is to create a stronger link between the participant and the outcome of the test. One can invite participants to translate their test results into concrete behaviours and thus justify the outcome of the test by giving clear examples of situations in which their behaviour matches their determined personality. Participants then identify behaviours at the desired endpoint and set a direction for growth. Obviously, further experimentation and research is necessary to develop this idea.

6 Conclusion and Future Work

In this paper we introduced a game-based workshop development format grounded in learning theory. The presented Artemis Bridge workshop, which is designed with this format, aims to further 21st century skills of aspiring professionals. Participants of the workshop were excited about the set up and execution and left with a proper idea how to improve their team performance.

The game-based workshop development format is a practical method. Although the Artemis Bridge workshop showed promising results, further development is required to better align individual parts of the proposed method. This serves two purposes. First, inexperienced facilitators will require instruments to support the subsequent stages of the workshop and second, aligning workshop parts will truly integrate multidisciplinary expertise of developers, learning theorists, psychologists, and engineers. Both purposes obviously contribute to the effectiveness of the workshop. A facilitator now typically runs one workshop with 6 participants at a time, with proper instruments a facilitator may run two of them in parallel. Our development format allows for substitution of parts, such as the MBTI and the Artemis Bridge game. In fact, our method dissects the learning process such that experts can focus on their expertise. Game developers should focus on the mechanics of the game to create team flow and engagement. Whereas engineers should focus on a measurement system and psychologists should focus on effective reflection, inference, and the measurement thereof. This in contrast to many past serious game projects where the responsibility of creating an engaging and serious game for a specific field or specific skill was predominantly the responsibility of the game development team.

Dissection of the learning process gives room for adjusting the workshop to individual needs. The goal to further 21st century skills might be similar for all aspiring professionals but is not identical. Their individual positions and the organisational context must be considered. Applying different games for different groups may turn out to be beneficial, such that the game would better suit individuals needs in interacting with the experience.

The MBTI test results show variations in groups of students. With a slight change in the setup of the workshop we may develop it into a training recommendation instrument for freshman students. To the least, we let them experience the personality of many of their fellow students in a particular field.

References

1. Artemis spaceship bridge simulator (2010). http://artemis.eochu.com/
2. Brandt, T.M., Edinger, P.: Transformational leadership in teams - the effects of a team leader's sex and personality. Gend. Manag. Int. J. **30**(1), 44–68 (2015)
3. Coovert, M.D., Winner, J., Bennett Jr., W., Howard, D.J.: Serious games are a serious tool for team research. Int. J. Serious Games **4**(1), 41–55 (2017)
4. Kaplan, R.M., Saccuzzo, D.P.: Psychological Testing: Principles, Applications, and Issues. Nelson Education, Toronto (2017)
5. Keith, M., Anderson, G., Dean, D.L., Gaskin, J.E.: The effects of team flow on performance: a video game experiment. In: SIGHCI Proceedings 13 (2014)
6. Kolb, A.Y., Kolb, D.A.: The learning way. Simul. Gaming **40**(3), 297–327 (2009)
7. Kolb, D.: Experiential Learning: Experience as the Source of Learning and Development. Prentice-Hall, NJ (1984)
8. Metfessel, N.S., Michael, W.B., Kirsner, D.A.: Instrumentation of bloom's and krathwohl's taxonomiesfor the writing of educational objectives. In: Proceeding of the AERA annual meeting (1969)
9. Phaneuf, J.É., Boudrias, J.S., Rousseau, V., Brunelle, É.: Personality and transformational leadership: the moderating effect of organizational context. Personal. Individ. Differ. **102**, 30–35 (2016)
10. Rosen, M.A., Salas, E., Wilson, K.A., King, H.B., Salisbury, M., Augenstein, J.S., Robinson, D.W., Birnbach, D.J.: Measuring team performance in simulation-based training: adopting best practices for healthcare. Simul. Healthc. **3**(1), 33–41 (2008)
11. Ross, S.M., Offermann, L.R.: Transformational leaders: measurement of personality attributes and work group performance. Personal. Soc. Psychol. Bull. **23**(10), 1078–1086 (1997)
12. Spath, D., Dworschak, B., Zaiser, H., Kremer, D.: Kompetenzentwicklung in der Industrie 4.0, Lehren und Lernen für die moderne Arbeitswelt, pp. 113–124 (2015)
13. Stewart, G.L.: A meta-analytic review of relationships between team design features and team performance. J. Manag. **32**(1), 29–55 (2006)
14. Thatcher, D.C.: Promoting learning through games and simulations. Simul. Gaming **21**(3), 262–273 (1990)
15. Wilson, K.A., Bedwell, W.L., Lazzara, E.H., Salas, E., Burke, C.S., Estock, J.L., Orvis, K.L., Conkey, C.: Relationships between game attributes and learning outcomes. Simul. Gaming **40**(2), 217–266 (2009)

Emotions Detection Through the Analysis of Physiological Information During Video Games Fruition

Marco Granato(✉), Davide Gadia, Dario Maggiorini, and Laura Anna Ripamonti

Department of Computer Science, University of Milan, via Comelico 39, 20135 Milan, Italy
{marco.granato,davide.gadia,dario.maggiorini,laura.ripamonti}@unimi.it

Abstract. Games are interactive tools able to arouse emotions in the user. This is particularly relevant in Serious Games, where the main goal could be educational, pedagogical, etc. Therefore, understanding the players' emotions during the game fruition could provide a valid support to the developers and researchers in video games field in order to design a more effective product. The presented research is a starting point to propose a framework for the determination of the player emotions through physiological information. We acquire several signals: facial electromyography, electrocardiogram, galvanic skin response, and respiration rate. We then compare the data to an emotional player assessment, defined using a valence and an arousal vector, through the application of Machine Learning techniques. The obtained results seem to suggest that the proposed approach can represent a valid tool to analyze the players' emotions.

Keywords: Video game · Serious game · Emotion detection
Machine Learning · Feature selection · Physiological data
Affective Computing

1 Introduction

Video games are interactive softwares, as Sid Meier says, structured by "a series of interesting choices" [1]. Thus, the players' decisions provide a specific gaming experience that, usually, may have a different emotional impact. In order to investigate the connection between decision making and emotions, Damasio developed the Somatic Marker Hypothesis (SMH) [2], the author described how the decisions are defined by previous outcomes. In addition, he described [3] how emotions alter physiological condition in relation to a specific stimulus, and how they can modify the future decisions. Starting from this concept, many researches

The original version of this chapter was revised: Modifications have made to Fig. 2. For detailed information please see erratum. The erratum to this publication is available online at https://doi.org/10.1007/978-3-319-71940-5_23

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 197–207, 2017.
https://doi.org/10.1007/978-3-319-71940-5_18

have investigated the relationship between video games and emotions [4–6]. A good video game can create engagement to the player, maximizing the emotions induced by the game choices. Thus, a game designer should consider and balance the different video game features in order to maintain players attention and to generate the desired emotional response [7]. Furthermore, a game has a subset of rules with which the players should interact in order to receive a game output (e.g. use a sword to kill an enemy). This mechanism lays the foundation to an *Affective Loop* [8] where the human-computer interaction becomes an emotional communication process.

A specific family of video games is composed by Serious Games. A serious game *"is a game in which education (in its various forms) is the primary goal, rather than entertainment"* [9]. Thus, serious games differ from the classic educational tools as they use a different framework, namely that of the game, to achieve an educational purpose using an entertaining product. Consequentially, as an entertaining and interactive software, a serious game should arouse specific emotions in order to engage the players and to transmit adequately the educational message as intended by the game designers.

In this paper, we propose a framework for the analysis of player emotions. We have collected physiological signals from a set of players during the video games fruition. These signals are: electromyography (EMG) on the players' face, in order to include the facial expressions in the analysis, electrocardiogram (ECG), with which we can acquire the information about RR variability, galvanic skin response (GSR), and respiration rate. We will use Machine Learning (ML) techniques (i.e. Random Forest RF [10]) in order to extract the important features from the signals, and to predict the self-assessment emotional data, provided by the user after the game session.

2 State of the Art

Video games can be used to: to *generate* [11] and *study* [12] player emotions. To generate emotions, several guidelines and notions have been proposed [11], like e.g., the "Magic Circle" [13], and the Lazzaro's four "fun" factor model [14]. These guidelines suggest how to develop and design a game in order to maintain the players' emotions at the center of players' experience [15]. If the player's emotions are well managed, the player can sense a feel of *Spatial Presence* [16] and *Flow* [17].

Spatial Presence and *Flow* are two theories that try to identify how a video game, or a generic entertainment product (e.g. a movie), interacts with the human emotions. *Spatial Presence* is a psychological condition where a player has the illusion to be transported in a virtual environment. Researchers suggested that a high sense of Spatial Presence can improve games entertainment and it may also facilitate the players performance [18]. The *Theory of Flow* is a theory of positive psychology that tries to define a mental state where a user is completely absorbed in a task. This theory describes a balanced channel between challenge and ability, where a person can profit an experience of achievement and happiness.

Emotions arise spontaneously in humans, and they may modify our decisions and actions. The theory proposed by James and Lange [19] states that the emotions are a reaction to physiological changes in arousal. Thus, physiologic reactions to emotions can be described by at least three human output systems [20]: self-report measure (e.g. through verbal expressions), behaviors (e.g. facial expressions), and physiological reactions of Autonomic Nervous System (*ANS*) e.g. Heart Rate, Brain Activity, etc. Hence, as emotions have a physiological response, they may be identified through a set of instruments able to detect the human physiology variations. This is the main argument at the basis of the Affective Computing research field [21], which suggest that any computer has the ability to express and recognize the people affects. In video games, Affective Computing is relevant in three main aspects: *players emotions recognition* - which leads to the game response to the emotions -, generation of '*affective behaviors*' in the game characters to enchant the realism with a credible output to various game events, and modelization of the *emotions' generations* on the game characters in order to represent a believable physical reaction (e.g. facial expressions) [22]. The possibilities of recognizing emotions, through physiological information, during the video games fruition, allowed the researchers and developers to investigate what kind of game events can generate a specific emotion and to design games aimed at inducing specific emotions in the players. Some interesting examples on this research field are [23,24].

2.1 Emotions Labeling

In literature there are two main theories of emotion labeling. Ekman [25] discretized six basic and universal human emotions: anger, fear, sadness, happiness, disgust, and surprise. These subgroups of emotions are universal and they have the same physiological outcome across the human cultures. Another methodology [26,27] is a dimensional model typically defined using two vectors: a *Valence* vector, that defines the emotion quality (from *averseness* to *attractiveness*), and an *Arousal* vector, that defines the emotion intensity (from *very calm* to *very excited*). In order to support the emotions recognition, a common practice is to use a set of Self-Assessment Manikins (SAM) [28] and to map them in the 2-dimensional space designed by the valence and arousal vectors. SAM can be considered as a scale usable to measure PAD (Pleasure/valence, Arousal, and Dominance). SAM uses a set of abstract characters (Manikins) defining a variable scale of values. In this paper, we will use only the information of valence and arousal; the value of dominance will be analyzed in future developments.

An additional tool to map the values of arousal and valence, developed by Betella and Berschure [29], is the Affective Slider (AS). It was designed in order to be closer to the users, using a contemporary symbolism, and to be comparable to the SAM.

3 Experiments

To investigate the players' emotions during video games fruition, we have performed a set of experiments with a group of users. We have asked the participants to play different video games. During the gaming sessions, we have recorded a set of physiological information, and we have also collected the emotional feedback using the self-assessment markers presented in Sect. 2.1.

3.1 Participants

The group of participants was composed by 10 males between 18 and 38 years old (in particular 50% of them have an age between 21 and 24 years). 90% of users usually plays video games more than 3 days by week and 30% plays every day. Furthermore, the 60% of participants declares that they play at home at least two hours for each game session. When asked about the preferred game genre, half of the participants have indicated platform games, also claiming to be sufficiently skilled in this genre. Unfortunately, due to technical issues, a participant was not eligible for the study, and then he was discarded.

3.2 Considered Games

All video games selected for the experiment are platform games. This category of video game is typically structured by a linear environment where the player starts from a position A and must reach a position B, avoiding obstacles and defeating enemies. Usually, platformers have simple mechanics (i.e. the character can jump, run, and attack the enemies), and a linear progression structured in levels.

This game category has received great changes in the game design and level design over the years (both in the game difficulty than in the interaction with the game environment). Thus, we had selected an heterogeneous topology of video games in this genre. The different indexes used to defines the topologies are: game environment (2d/3d), release date (recent/"classic" games), and license (commercial/non-commercial games). Anyway, almost all games have received a good rating by the specialized critic or, in any case, they are considered *enough fun to play* [30,31]. The selected games are:

- *Rayman Origins* (2011) is the fourth chapter of its series and it is developed by *Ubisoft*. It consists in a 2D side-scrolling platformer with collectible items. The players have to play the first level (as a tutorial) and the third level in the third stage.
- Six platform levels generated with *FunPledge 2.0* (developed in 2016)[31]. Evolution of a previous work [30], it is a tool able to generate 2D side-scrolling platform levels automatically, on the basis of a musical rhythm used as a basis to define challenging and entertaining levels. The players should finish all the levels. The game session was also stopped if the player reaches the game over.
- *Earthworm Jim 2* (1995) is the second chapter of a series developed by *Shiny Entertainment*. It is a 2D side-scrolling platform where the protagonist can also shoot to the enemies. The player should finish the first level. The game session was also stopped if the player reaches the game over.

– *Crash Bandicoot* (1996) is the first game of a series initially developed by *Naughty Dog*. It is structured by 2D side-scrolling levels and 3D levels. The players have to play the thirteenth level (a 3D level). The game session was stopped also if the player reaches two game overs. This softening of end-of-session rules is given to allow the user to switch between an analog cursor to the directional crosses.

3.3 Devices and Physiological Signals

From each player, we acquired different physiological data: electrocardiography (ECG), electromyography (EMG) on 4 facial muscles, galvanic skins response (GSR), and respiration rate. The data was collected with a set of sensors connected to an Arduino Due, that performs a 12-bit quantization with its ADC converter and it sends the data to a computer through Serial Communication (virtualized on USB). For the ECG and EMG sensors, we used an Olimex-EKG-EMG shield. It is a device already used in biomedical engineering [32] and it is able to read a 3-lead electrode connector via 3.5 jack.

In order to acquire the ECG, we connected three electrodes, two at both wrists and one at the left ankle, following the Einthoven's triangle [33]. The EMG sensors were placed on 4 different areas of the right side of the players' face as illustrated in Fig. 1: on *Zygomaticus Major* (*EMG1*), *Orbicularis* (*EMG2*), *Nasalis* (*EMG3*), and *Supercilli* (*EMG4*) muscles. Lastly, each EMG sensor uses as reference a common electrode, placed on the hair border, as suggested in [34]. The GSR data were collected placing two electrodes on two different distal phalanxes (index and middle fingers) of the left hand not used to control the selected games, and the electric voltage was amplified using a LM324 Surface Mount Device (SMD) installed on a Grove GSR sensor. GSR can be considered as a reflection of the sympathetic axis that produces an eccrine sweat gland [35]. In humans the sweating has also a function of emotional expression [36]. Thus, sympathetic activity is linked to the emotions and, therefore, GSR is often suggested as an emotions index [37]. The respiration intensity/rate was measured placing a digital thermometer (DS18B20) under the player's nose. The thermometer has been placed avoiding contact with the player's skin, limiting, as a consequence, the noise involved with the epidermal temperature. The data were stored with an 11-bit resolution (it has a fluctuation of 0.125 °C) and sampled each 375 ms as indicated by the DS18B20 Datasheet [38].

3.4 Procedure

The participants were invited to sit on a comfortable chair. They were informed about the experimental procedure and they were invited to read and sign an informed consent, and a permission to use the video recorded during the experiment for research and academic purposes. The acquired data was collected anonymously, and at each participant was assigned an incremental identifier to be used for the analysis and for future experiments. Lastly, each user performs the experiment in a unique daily session. During the experiment, a 5 MegaPixel,

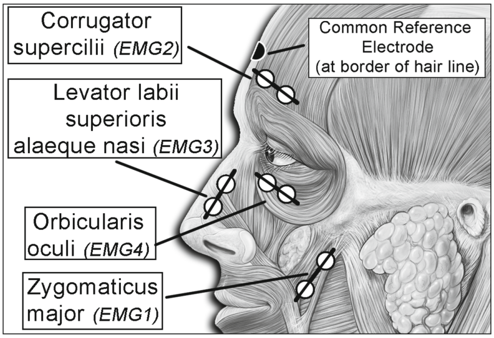


Fig. 1. Representation of facial electrodes positions connected to EMG sensors. Original medical illustration from Patrick J. Lynch (https://goo.gl/ttgxo6)

1280×1024, camera was placed on the 32″ display used during the gaming sessions. The games were played using a gamepad on a computer with Windows 10 OS, i7 6700k CPU, 32 gb RAM (DDR4), and NVIDIA GeForce GTX 1080 GPU.

The experiment consisted of three different stages: electrodes placement, game session, and emotion labeling. In the first stage, the sensor were placed taking care to not bother the players during the session. In the second stage, the participants should play all the levels, with the order as described in Sect. 3.2. Before each game session, the games mechanics and goal (as described in Sect. 3.2) were explained to the participant. Before and after each level, the player had to press an analogical button (connected to the Arduino) in order to synchronize the physiological data with the game sessions. In the third stage, the participants were asked to map their emotional states on the arousal/valence dimensional model (see Sect. 2.1). A video of the recorded game session and of the participant's face during the playing time was shown to the player. After a training, the player uses the two gamepad analog cursors to set two pointers, one for arousal and one for valence, following the SAM and the AS markers described in Sect. 2.1. She could also rewind the video in order to correct her choice evaluation. Lastly, each user was asked to answer to a survey with questions regarding her gamer skills and habits, and regarding the overall experiment considerations.

4 Data Analysis

The physiological data were collected through an Arduino with a variable sample rate SR between 952 Hz and 989 Hz, due to the time fluctuation of the temperature sensor (see 3.3). Thus, each signal second was sub-sampled to the lowest value. Moreover, the data about emotional labeling were sampled at the same frame rate of the video (30 samples/second).

The signal of each physiological information was smoothed using Savitzky/Golay filter [39], applied with a polynomials fit of order 30 and a floating window with a length equal to the odd value nearer to half sample rate (i.e. 475).

This filter is typically used to smooth noisy signals, such as electrical ones. The filtered signals were also de-noised using an orthogonal wavelet with 5 levels of decomposition. A *Penalized Contrast Function* [40] was used to identify the location and the variance change points.

In order to acquire the players' Heart Rate (HR) during the game session, we have detected the QRS complex (where Q is a downward deflection wave, R is a upward deflection wave, and S is any downward deflection wave after the R wave) on the ECG signal using OSEA algorithm [41]. From the QRS data, we get only the information about the position of R waves which, in ECG, represents the left apex depolarization, in order to obtain RR variability, and, consequently, the HR variability. A second analysis of the ECG was performed to find the motion artifacts (or signal zones where the QRS can not be detected), and, in order to have a real approximation of RR intervals, in these areas it was applied an interpolation of RR peaks.

Then, each physiological signal (except for the ECG) was sampled with a range of half second and for each subsample the values between 1 Hz to 180 Hz in the frequency domain (using FFT) were calculated. Lastly, the self-assessment data (30 data/s) was interpolated in order to have the same length of the other signals (2 data/s). Summarizing, the features (583) used for predicting the arousal and valence values are: row data of each physiological signal (avoiding ECG), HR, and magnitude information of each physiological signal considering only the frequencies between 1 Hz to 180 Hz.

In order to identify the most informative features, we randomly split our sample in two different groups called TRs and TE. TRs is composed by 8 participants (a matrix with 40954 instances and 583 features) used to train the ML algorithm; it was used to determine the most informative features. TE group is composed by the remaining participant (a matrix with 6510 instances and 583 features), and it was used to test the features extracted by the ML trained on TRs. In order to have a robust feature selection, we randomly created 50 different permutations of a subset by 6 subjects in TRs. The data of these participants were merged and we train a RF on them in order to predict the self-assessment data of valence and arousal, labeled by the users as described in Sect. 3.4. The RF creates a bagging of decision trees [10], and it is able to extract an index of importance for each features as described in [42]. After each training, the importance of each feature was stored obtaining a matrix $M = RxC = 50x583$. For each row, we extracted the most important k features, where k is an incremental iterator from 1 to C. Among the remaining features, we selected those available in all permutations (R), and we extracted the features from the joined data of TRs. With these data, we trained a RF using 1/3 of features for each decision split and with 128 trees as suggested in [43]. It was used to predict values of the user in the saving the Root Mean Square Error (RMSE) index calculated between the prediction data and the self-assessment data. Thus, after all interactions, we had two sets of C RMSE values, one for arousal values and one for valence values. On these vectors, we had selected the features that minimize the RMSE value. A table with the most important features is available at the url https://github.com/grano00/EmotionsAnalysisInVideogames.

5 Results and Discussions

The most important features were used to train two RFs in order to predict the data on TE:

– A general predictor, using a RF trained on the joined data of the 8 participants TRs (i.e. Leave One Out validation).
– An individualized predictor, using a RF trained using a ten cross validation [44] on the TE.

Figure 2 illustrates the predictions data trends on the self-assessment values during the TE game session. The plots 1 and 2 are the predictions trends on the data trained on TRs. Albeit the RMSE values (1.6 for arousal and 1.2 for valence) do not indicate a statistical significance of the results, we can however observe from the graphical plots that the values remain in the emotional range provided by the user TE. Predictions represented in plots 3 and 4, trained with the cross validation on TE, follow coherently the data labeled by the user TE through self-assessment (Sect. 3.4). Also in this case, due to the users variability in the self-assessment methods, with the consequent introduction of several peaks in the evaluation data, the observed RMSE index does not provide a significant information. Although the prediction data does not follow the peaks in

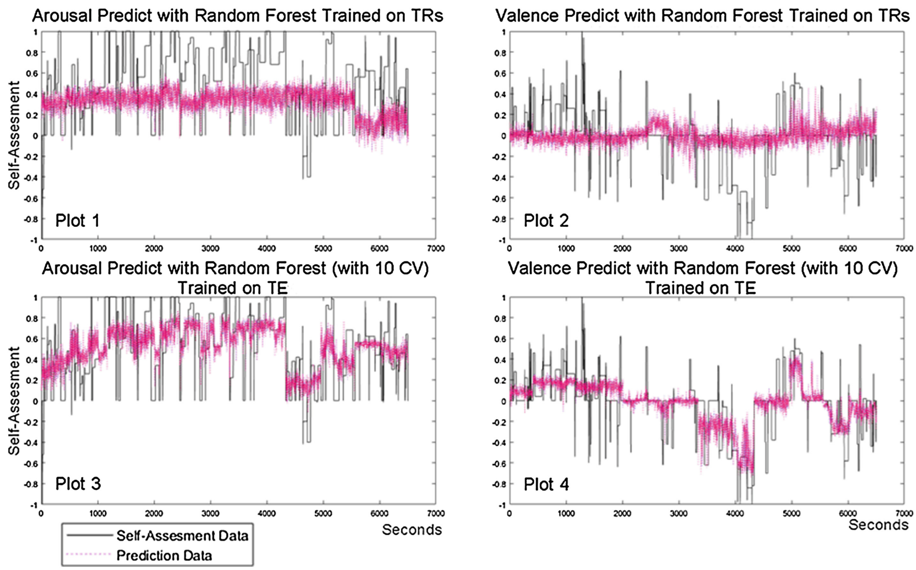


Fig. 2. The figure represents the mapped value of TE on emotional assessment. The time of each game session (using half second as step unit) is shown on x axis, the emotional assessment values are shown on y axis. The prediction values of the first two plots are defined by a Random Forest trained on TRs, while, the remaining two plots illustrate the predictions of a Random Forest trained with a 10 cross validation on TE.

the self-assessment values, they follow the pattern identified by the user TE. The prediction of the cross-validation analysis on TE is more accurate than the prediction of RF trained on TRs, since the self-assessment data is user sensitive. Anyway, this research is a work in progress and we hypothesize that, collecting additional data on a larger sample of users, the difference between the two predictions can be attenuated.

In conclusion, we consider the described approach as a potentially valid starting point for the developing of a framework to analyze the players' emotions. It could be used to understand which part of a game can generate a specific physiological state, and, therefore, a corresponding emotion. As said in Sect. 1, the main purpose of a serious game is to educate the users. The emotions are related to learning [45], thus a similar framework could help the evaluation about the effectiveness of serious games. For example, during the beta test of a game, the developers can identify the game areas which unwittingly induce unwanted emotions, and, consequentially, redesign them. In addition, the developers can handle and design an algorithm for a real-time adaptation of some game mechanics (e.g. the mechanics that affect the difficulty), in order to avoid specific unwanted emotions (e.g. boredom). These techniques may help the players' engagement, and, therefore, the possibilities to transmit a the desired message.

Future works will focus on a full validation with a larger number of subjects. This extension could improve the accuracy of the data analysis and, consequentially, the prediction outcome. Future extensions will also include:

– Extension of the experiment using different game categories (e.g. serious games),
– selects interesting parts of the game session and ask to external experts to re-label the player emotions,
– improve the filtering on the physiological signals, e.g. filter the natural increment of the GSR over the time,
– analysis of the players' emotions in a discrete domain, in order to understand the player fluctuation over the flow channel.

References

1. Fullerton, T.: Game Design Workshop: A Playcentric Approach to Creating Innovative Games. CRC Press, Boca Raton (2014)
2. Damasio, A.R., Tranel, D., Damasio, H.C.: Behavior: theory and preliminary testing. In: Frontal Lobe Function and Dysfunction, p. 217 (1991)
3. Damásio, A.R.: Descartes' Error: Emotion, Reason, and the Human Brain. Quill (1994)
4. Grodal, T.: Video games and the pleasures of control. In: Media Entertainment: The Psychology of its Appeal, pp. 197–213 (2000)
5. Granic, I., Lobel, A., Engels, R.C.M.E.: The benefits of playing video games. Am. Psychol. **69**(1), 66 (2014)
6. Isbister, K.: How Games Move Us: Emotion by Design. MIT Press, Cambridge (2016)

7. Bateman, C.: Beyond gAme Design: Nine Steps Toward Creating Better Videogames. Cengage Learning (2009)
8. Höök, K.: Affective loop experiences – what are they? In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 1–12. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-68504-3_1
9. Michael, D.R., Chen, S.L.: Serious Games: Games that eDucate, Train, and Inform. Muska & Lipman/Premier-Trade, Cengage (2005)
10. Ho, T.K.: Random decision forests. In: Proceedings of the Third International Conference on Document Analysis and Recognition, vol. 1, pp. 278–282. IEEE (1995)
11. Koster, R.: Theory of Fun for Game Design. O'Reilly Media Inc., Sebastopol (2013)
12. Ravaja, N., Turpeinen, M., Saari, T., Puttonen, S., Keltikangas-Järvinen, L.: The psychophysiology of James Bond: phasic emotional responses to violent video game events. Emotion **8**(1), 114 (2008)
13. Salen, K., Zimmerman, E.: Rules of Play: Game Design Fundamentals. MIT Press, Cambridge (2004)
14. Lazzaro, N.: Why we play games: four keys to more emotion without story (2004). http://www.xeodesign.com/whyweplaygames/xeodesign_whyweplaygames.pdf
15. Yannakakis, G.N., Paiva, A.: Emotion in games. In: Handbook on Affective Computing, pp. 459–471 (2014)
16. Draper, V.J., Kaber, D.B., Usher, J.M.: Telepresence. Hum. Factors **40**(3), 354–375 (1998)
17. Nakamura, J., Csikszentmihalyi, M.: The concept of flow. Flow and the Foundations of Positive Psychology, pp. 239–263. Springer, Dordrecht (2014). https://doi.org/10.1007/978-94-017-9088-8_16
18. Ravaja, N., Salminen, M., Holopainen, J., Saari, T., Laarni, J., Järvinen, A.: Emotional response patterns and sense of presence during video games: potential criterion variables for game design. In: Proceedings of the Third Nordic Conference on Human-Computer Interaction, pp. 339–347. ACM (2004)
19. James, W.: What is an emotion? Mind **9**(34), 188–205 (1884)
20. Mauss, B.I., Robinson, M.D.: Measures of emotion: a review. Cogn. Emot. **23**(2), 209–237 (2009)
21. Picard, R.W.: Affective computing (1995)
22. Hudlicka, E.: Affective computing for game design. In: Proceedings of the 4th International North American Conference on Intelligent Games and Simulation (GAMEON-NA) (2008)
23. Nacke, L.E.: An introduction to physiological player metrics for evaluating games. In: Seif El-Nasr, M., Drachen, A., Canossa, A. (eds.) Game Analytics, pp. 585–619. Springer, London (2013). https://doi.org/10.1007/978-1-4471-4769-5_26
24. Ninaus, M., Kober, S.E., Friedrich, E.V.C., Dunwell, I., De Freitas, S., Arnab, S., Ott, M., Kravcik, M., Lim, T., Louchart, S., et al.: Neurophysiological methods for monitoring brain activity in serious games and virtual environments: a review. Int. J. Technol. Enhanced Learn. **6**(1), 78–103 (2014)
25. Ekman, P.: An argument for basic emotions. Cogn. Emot. **6**(3–4), 169–200 (1992)
26. Bradley, M.M., Greenwald, M.K., Petry, M.C., Lang, P.J.: Remembering pictures: pleasure and arousal in memory. J. Exp. Psychol. Learn. Mem. Cogn. **18**(2), 379–390 (1992)
27. Lang, P.J.: The emotion probe: studies of motivation and attention. Am. Psychol. **50**(5), 372 (1995)

28. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry **25**(1), 49–59 (1994)
29. Betella, A., Verschure, P.F.M.J.: The affective slider: a digital self-assessment scale for the measurement of human emotions. PLoS ONE **11**(2), e0148037 (2016)
30. Ripamonti, L.A., Mannalà, M., Gadia, D., Maggiorini, D.: Procedural content generation for platformers: designing and testing FUN PLEdGE. In: Multimedia Tools and Applications, pp. 1–50 (2016)
31. Mazza, C., Ripamonti, L.A., Maggiorini, D., Gadia, D.: Fun pledge 2.0: a funny platformers levels generator (rhythm based). In: To be presented at CHItaly 2017: the Biannual Conference of the Italian SIGCHI Chapter, Cagliari, Italy, September 2017 (2017)
32. Stojanović, R., Čaplánová, A., Kovačević, Ž., Nemanja, N., Bundalo, Z.: Alternative approach to addressing infrastructure needs in biomedical engineering programs (case of emerging economies). Folia Medica Facultatis Medicinae Universitatis Saraeviensis, **50**(1) (2015)
33. Conover, M.B.: Understanding Electrocardiography. Elsevier Health Sciences (2003)
34. Van Boxtel, A.: Facial EMG as a tool for inferring affective states. In: Proceedings of measuring behavior, pp. 104–108. Noldus Information Technology Wageningen (2010)
35. Critchley, H.D.: Electrodermal responses: what happens in the brain. Neuroscientist **8**(2), 132–142 (2002)
36. Darwin, C., Ekman, P., Prodger, P.: The Expression of the Emotions in Man and Animals. Oxford University Press, USA (1998)
37. Venables, P.H., Christie, M.J.: Electrodermal activity. Tech. Psychophysiol. **54**(3), 3–67 (1980)
38. Ds18b20 datasheet. https://datasheets.maximintegrated.com/en/ds/DS18B20.pdf. Accessed 21 July 2017
39. Savitzky, A., Golay, M.J.E.: Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. **36**(8), 1627–1639 (1964)
40. Lavielle, M.: Detection of multiple changes in a sequence of dependent variables. Stochast. Processes Appl. **83**(1), 79–102 (1999)
41. Hamilton, P.: Open source ECG analysis. In: Computers in Cardiology, pp. 101–104. IEEE (2002)
42. Loh, W.-Y.: Regression tress with unbiased variable selection and interaction detection. Statistica Sinica **12**, 361–386 (2002)
43. Bonab, H.R., Can, F.: A theoretical framework on the ideal number of classifiers for online ensembles in data streams. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, pp. 2053–2056. ACM (2016)
44. Kohavi, R., et al.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: IJCAI, Stanford, CA, vol. 14, pp. 1137–1145 (1995)
45. Csikszentmihalyi, M.: Flow and the Psychology of Discovery and Invention. Harper Collins, New York (1996)

Design of a Component-Based Mobile Learning Game Authoring Tool

Pierre-Yves Gicquel(✉), Sebastien George, Pierre Laforcade, and Iza Marfisi-Schottman

Laboratoire d'Informatique de l'Université du Maine (LIUM - EA 4023), Le Mans Cedex 9, France
{pierre-yves.gicquel,sebastien.george,pierre.laforcade, iza.marfisi}@univ-lemans.fr

Abstract. Mobile learning games (MLGs) have great potential in education, especially in fields requiring outdoors activities such as botany or cultural heritage education. However, the number of mobile learning games actually used for outdoor education remains very low. The absence of dedicated applications allowing to build MLGs, without technical expertise, is certainly one of the most important factors. To overcome this limitation, we propose a design method for MLG authoring tools, based on reusable components. We describe in this paper the design and first results of Moggle-Designer, a full-web MLG authoring tool requiring no technical expertise. We detail how a model-based approach, combined with component-oriented programming, adequately allow the representation and manipulation of didactic expert knowledge. Finally, we present Moggle-Player, an application for running the designed MLGs on any mobile browser. Moggle-Designer was used to design several mobile games that have been tested during a pilot study in a botanical park.

Keywords: Mobile learning · Serious games · Authoring tool
Situated learning

1 Introduction

Mobile devices, such as mobile phones or tablets, have proven to be interesting mediums for computer supported learning. These devices are particularly interesting out of the classroom, for authentic learning activities. Mobile Learning Games (MLGs) present great pedagogical interest because they favor environment exploration and self-regulated mediation through conversation between learners [1]. These exploratory and conversational activities allow sustainable learning and appropriation through collaborative knowledge building [2].

However, the cost for producing usable MLGs remains prohibitive for most educational structures. The few available authoring tools are either very generic, and require a strong technical expertise, or only allow the design of overly simple and specific learning scenarios [3].

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 208–217, 2017.
https://doi.org/10.1007/978-3-319-71940-5_19

This raises the question of how to design a simple and generic MLG authoring tool. In this paper, we propose the use of a component-based approach. First, we define our research context and questions, then we present the methodology for collecting the needs of teachers in terms of games scenarios and for formalizing these scenarios through a model driven approach. We will then present the Moggle-Designer authoring tool, along with Moggle-Player, which allows users to play the authored MLGs on any mobile browser. Finally, we present the results of a first experimentation in real conditions, in which subjects played various gaming scenarios conceived with Moggle-Designer in a botanical park.

2 Research Context and Questions

The research context of this paper is the ReVeRIES project[1] (French acronym that translates to "dreams" and that stands for Educational, Interactive and Fun Plant Recognition on Smartphones). This project aims to introduce mobile applications in botanical education. It relies on the assumption that the urbanization of society has drastically reduced common botanical knowledge, leading to a gradual increase in so-called *plant blindness*, defined by [4] as *the inability to recognize the importance of plants in the biosphere and in human affairs*.

In order to motivate children and botanical neophytes to learn about plants and explore their natural environment, we intend to design MLGs that will help to learn about plant characteristics. We define a MLG as a mobile application, combining educational content with a playful scenario to favor learning. This kind of application relies on mobile devices' features (e.g. location, orientation and proximity sensors, media capturing and recording, augmented reality...) to contextualize activities. One of the project goals is to propose an MLG authoring tool that does not require any programming knowledge. This authoring tool has a generic base but specific features could be added for enriching MLG with domain-specific activities. In the case of botany, we propose activities based on automatic image analysis to recognize plants from photos of their organs (e.g. leaves, barks, flowers). This aspect is detailed in [5].

Conceiving an authoring tool in the context of mobile learning games raises research questions we address in this paper:

- How to collect and formalize designers' needs, expressed in natural language, in an appropriate formalism for conception? (Sect. 3)
- How to facilitate the creation of MLGs based on this formalism and deploy them on the players smartphones? (Sect. 4)

3 Modeling Mobile Learning Games

This section is dedicated to the modeling issues related to the complexity of the MLG authoring tool. Indeed, the authoring tool should allow the design

[1] http://reveries-project.fr.

of a large variety of MLGs. Such learning scenarios require a domain-specific modeling language, at least as a conceptual language to drive the design and development of the authoring-tool. Designing a generic but still botany-oriented modeling language relevant for any botanical and mobile learning scenario is not an achievable objective. Our modeling approach aims to capture, explicit and validate a broadly-usable instructional language. This language being then extended with botany specific concept for the design of botany MLGs scenarios.

The next sub-sections detail our global approach, the meta-model we propose, and some feedback about its validation for designing botanical MLG scenarios.

3.1 Overview of the Modeling Process

Our modeling process follows several steps:

1. **Collecting informal (textual, drawings...) scenarios of botanical MLGs:** collaborative sessions with various actors involved in the ReVeRIES project (including botanical park managers, botanical scientists, and teachers) were conducted. Participants were split into groups and had to propose a botanical MLG for the *Echologia* natural park in Laval, France (a map was available about the tree species present in the park). The main objective was to identify and collect a large range of MLG activities. Six scenarios were produced and discussed during a collective retrospective [6].
2. **Describing a conceptual model for botanical MLG scenarios:** it offers high-level concepts, similar to the natural concepts used by teachers, and is rich enough to adapt to various types of MLGs [7]. Without going into detail, the model formalizes a MLG as a series of *SituatedGameUnits* each related to a *POI (Point Of Interest)*, i.e. a geographical zone involving zero or several *objects of interest* (mainly plants although not restricted to them). Each *SituatedGameUnits* can be described in terms of optional elements: *triggers* (to start the unit), *clues* (to help the player find the POI), *guidance* features to the POI (marker on a map, GPS, vibrations, etc.), *validation* actions (to validate the presence of the player on the POI), *OnSiteActivities* (activities to do around the POI), *rewards* (points/items/... in relation to the arrival on the POI and/or according to the success of the OnSiteActivities), *pedagogical information* (to debrief with the player the knowledge acquired in relation to this unit).
3. **Formalizing a specific meta-model for botanical MLG scenarios**.

Because the first two points have already been detailed in previous publications ([6,7]), we detail here the third point: the botanical MLG meta-model.

3.2 Meta-Modeling Approach

The meta-model we propose is a formal and explicit description of all the concepts, properties and relations required to model a botanical MLG scenario. MLG scenarios are the models, and they are conform to this meta-model. This meta-model has several objectives:

- to verify that the MLG scenarios can be formally specified (i.e. machine-readable with no ambiguity) in order to be exploited by the MLG player;
- to drive the design of the MLG authoring tool: we can check the consistency between the scenario elements (i.e. the visual information depicted in the mock-up screens and the interactions between screens);
- to drive the specifications and the development of the components-based MLG player engine.

One of the main issues to tackle is to ensure the expressiveness, completeness, consistency and the non-reducibility of the meta-model. This will help reduce the design and development complexity of the authoring and player tools. In order to obtain such a meta-model, we progressively merged all the models of MLGs it needs to cover. Indeed, during the creativity session with the ReVeRIES members, we realized that the designed MLGs actually had three different types of scenarios:

- linear scenario (e.g. treasure hunt): game units are chained up (use of guidances to drive the player to the POI);
- emergent scenario (e.g. interactive walk): game units start when players are physically close to a POI;
- hub scenario (e.g. geocaching): all game units and clues are available and depicted on a map, players decide in which order they want to play them.

For each of these scenario types, we therefore proposed a dedicated meta-model, and merged it to the meta-model. The merge relies on a (meta-)model composition in the meaning of the Model-Driven Engineering (MDE) principles. Briefly, we concretely merge the meta-concepts sharing a same semantics, gathering their potential different meta-attributes and meta-relations. Concepts that cannot be aligned are added, while elements that can be mapped to existent ones are removed. We also integrated game ingredients (e.g. rewards, points, inventories and virtual items) in order to cover all the elements of the designed MLGs.

Concretely speaking, this meta-model was specified using the Eclipse Modeling Framework [8] included in the *Obeo designer community* IDE. Specifying the MLG meta-model allowed to describe formally the cardinality and restriction between the elements of the meta-model (such as *OnSiteActivities*).

4 Moggle-Designer

4.1 Presentation

Moggle-Designer (MObile Geolocated Games for Learning) allows non-programmers to create MLGs with a bottom-up design process: learning and gaming components, such as media (e.g. images, videos, formatted text) or point of interest, are first defined as basic building blocks.

These blocks can then be aggregated to construct more complex components. For instance, a guidance activity is composed of a point of interest (the destination) and a medium (feedback) that is displayed when the user reaches this destination. Finally, these building blocks are linked together to create a game unit, which can be seen as a self-contained, independent and complete MLG [7].

Game Units. Game units follows a similar pattern, organized in three phases. This first phase is guiding the player to the place of interest (POI). The designer can chose between two guidance modes, either the POI and user position are displayed on a classic map on the player phone, or a more playful mode where a circle grow as the user get closer to the POI.

Second phase is triggered by the player reaching the POI. The player is then proposed a learning task where she has either to answer questions related to the plant at this POI, or to correctly identify a specie by scanning the correct QR-code (various QR-code being placed on three at the POI). When the player correctly complete a learning task, she is rewarded by a bonus item stored in a personal inventory. For instance the bonus "recipe of rozenbottel jam" is acquired when the user correctly identify a Rosa Canina.

The third phase consist of proposing a feedback media that gives complementary information about the learning task (e.g. information such as "an easy way of identifying Rosa Canina from its fruit color").

4.2 Design Principles

Reusable and Shareable Building Blocks. One common limitation of authoring tools in Educational Game Authoring tools, is the lack of non-trivial working examples that teachers can use/reuse to guide them [9]. To reduce this issue and favor appropriation of Moggle-Designer by instructors, it encourage, through its design, sharing and reusing game components at different levels of granularity (from simple media component to game unit and complete scenario).

Figure 1 illustrates the relation of composition between shareable building blocks. When an instructor uploads an image file, it appears as a shareable component. The instructor can then create a document containing the image accompanied by informative text. This newly created document, is then used as part of a Multiple Choice Question.

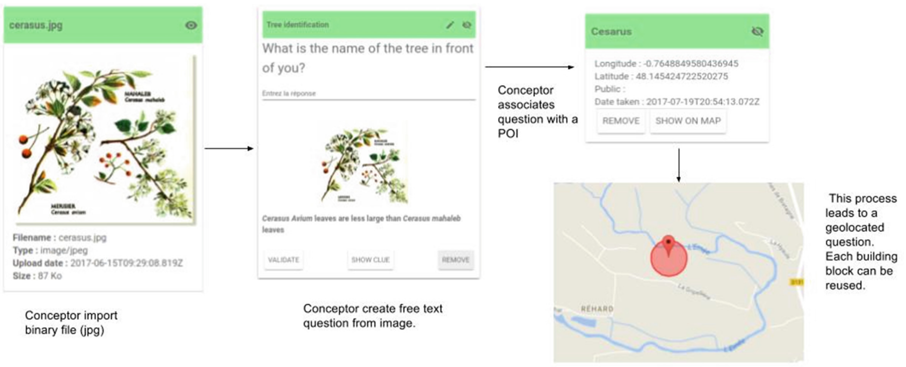


Fig. 1. Creating a situated question by combining building blocks

Ownership and Security of Building Block. In order to favor sharing while keeping a personal space for the user, we designed the system such as unidentified users could not access any building blocks. Moreover, users have the liberty to easily share or unshare each of their own contributions. Each block has a 1-1 relation with its creator's user-id: the blocks being stored in a document-oriented database, are stored, as a whole, with each of their properties. This allows to easily create unique hash from blocks and track their use.

Terminology and Usability. [10] underlines the critical importance of clear terminology and guidance in a user interface which should *speak the user's language*. In the case of an authoring tool using building blocks, these block must be easily apprehended by the user.

By proposing different granularities in the building blocks, with complex blocks being created from more simple blocks, we aim at easing the learning curve. An unexperienced user will easily be able to apprehend the principles of simple building blocks (e.g. image and text). Having acquired confidence and effectiveness in the conception of simple building block, the process of creating more complex building block by composition of simple ones will be more intuitive. Moreover, in order to encourage designers to explore and experiment, a direct visual feedback (WYSIWYG) is displayed during the block creation.

Advantages and Technical Challenges of Web Applications. We decided to build Moggle-Designer as a web application. Various reasons guided this choice: first, web applications are platform independent. Moreover, the distribution of a web application is straightforward, there is no installation process for the users and the URL always point to the last stable version. From a technical point of view, we used the Web Component paradigm[2], a recent evolution in Web Application promoted by the W3C. The fundamental idea behind Web Components is to create easily reusable pieces of HTML and JavaScript code. Web Components provide a way to encapsulate the HTML/CSS and encapsulate the JavaScript within the context of execution of the component. This means one can use someone else's components with the complete guarantee that there will be not be any side effects on his/her own components.

4.3 Moggle-Player

Moggle-Player is a tool that allows users to play scenarios designed with Moggle-Designer. It is also a full web application, thus it is usable on any mobile device with a browser (Iphone, Android, Windows phone...). The game player makes full use of mobile HTML5 capacities (Fig. 2): use of GPS for guiding the user to the different games steps, use of the phone camera for activities based on scanning a QR-code to identify a tree and use of websockets to permit chat between users.

[2] https://www.webcomponents.org/specs.

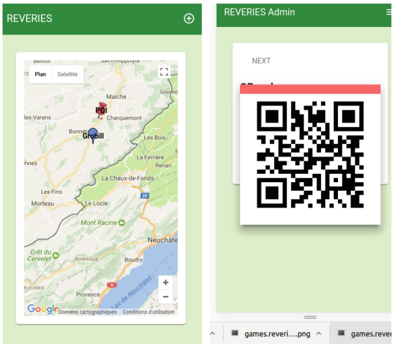


Fig. 2. Use of GPS and phone camera

One particular challenge we faced was to keep the web application running in absence of Internet connection. We set up a cache system (a JavaScript service worker) that preloads everything resources needed to keep the current MLG running when there is a disconnection and synchronizes back with the server when the network connection is restored.

5 Qualitative Pilot Study

We conducted a pilot study in a botanical park to evaluate how MLGs designed with Moggle-Designer were used and how the activities were perceived. We used three of the MLG scenarios designed during the ReVeRIES design sessions.

5.1 Participants and Apparatus

Twelve users took part in the experimentation (four on each MLG scenario) but only nine completed the feedback form, because some users actually shared the same smartphone. The nine that answered the feedback form were aged between 18 and 38 years old (mean: 28, SD: 8), and included four women and five men. Users where from three different backgrounds: four where fellow researchers from our laboratory, five were students in first year of computer science studies and three were working at the botanical park. User's personal smartphones where used for the experiment (these included iPhones and Android devices). All the devices were connected to a wireless mobile network, with variable connectivity.

5.2 Organization

We voluntarily gave little information to the players in order to assess the clarity of the MLGs. Three animators from the project team were present in case of technical problems, but they were actually solicited very rarely. The main reasons of solicitations were problem concerning geolocation, some smartphones having poor GPS accuracy.

The sessions lasted approximately 45 min. The MLGs were composed of six situated unit games each, implying a short walk through the park to reach the points of interest and progress in the game.

5.3 Main Results

We collected direct verbal feedback during and after the activities and also asked users to complete a questionnaire to assess their satisfaction and their interest in MLGs in the context of a botanical park. The verbal feedback was very positive: users enjoyed the fun aspect of having to reach a specific zone and having to observe the plants in the surroundings in order to answer questions.

The analysis of the questionnaire confirmed the verbal feedback: user found the MLGs motivating for the visit and useful to acquire knowledge in that botany. The questionnaire consisted mostly of Likert based unique choices questions with four possible items. A synthesis of the answers is presented in the Fig. 3.

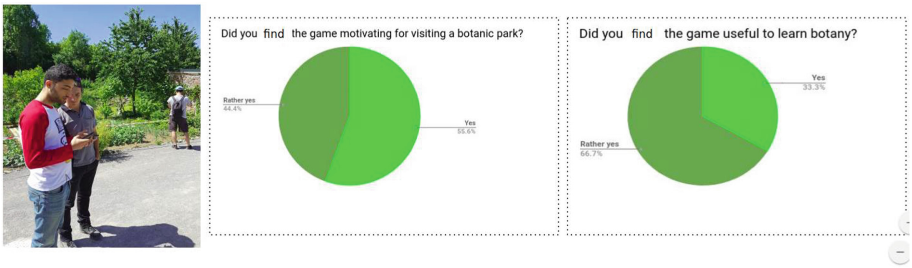


Fig. 3. Results of feedback form

While encouraging, these results need to be confirmed in a larger scale study. Nevertheless, this study demonstrates that Moggle-Designer can produce complete and usable MLGs.

6 Related Works

In this section, we compare two well established authoring tools for mobile learning applications to Moggle-Designer. A complete state of the art can be found in [11].

6.1 ARIS

ARIS is a Web based authoring tool that generates IOS Mobile Learning Games. While generating native mobile applications has some benefits, in terms of performance for instance, it greatly limits the possible users. Nevertheless, ARIS is particularly powerful, it allows the creation of different kinds of MLGs, from treasure hunts to location-based storytelling [12]. However, this expressive power comes with a steep learning curve. [13] argues about ARIS that the *"challenge is learning the user interface and game logic [...] many who admire the possibilities of this platform find creating and understanding the logic of requirements for objects in a game scenario challenging."* Finally, one key difference between ARIS authoring tool and Moggle-Designer is the inability to share resources in ARIS. This design strongly limits re-usability of learning objects.

6.2 MLearn4Web

MLearn4Web [14] is a full web authoring tool allowing teacher to create web learning application. The tool is very easy to apprehend, applications are built as a series of pages, each page being constructed visually by drag-and-dropping items on the page from a component pallet. First study shows a very good degree of acceptance among teachers.

However, MLearn4Web presents limitations in terms of design possibilities. Applications can only be designed and shared as a whole, that is to say it's not possible to design subpart of an application or to reuse part of someone else application. Moogle design principles in contrast allows sharing at different level of granularity, which makes produced games modular and easily reusable. Moreover, MLearn4Web do not allow to implement game mechanics, such as challenges or rewards for instance. This limit the learner experience to be more passive than active. In contrast, the application designed with Moogle incorporate game mechanism, requiring the learner to actively investigate her environment to answer questions.

7 Conclusion and Future Work

We present, in this paper, a complete design process for MLG (Mobile Learning Game) authoring tools, from the collection of users' needs to implementation details. The MLG scenarios were formalized through a model-based approach while web-components structurally and semantically conform to the model elements were implemented.

This process led to **Moggle-Designer**, an authoring tool powerful enough to design the types of scenarios expressed by future users, while needing no technical expertise to build these scenarios, and **Moggle-Player**, capable of playing these MLGs on any mobile device. An experimental validation in the real setting of a botanical park allowed us to validate the expressiveness of the authoring tool. We were able to implement scenarios that were proposed during the collection of users' needs phase. We validated that these scenarios were playable and first results suggested a strong interest and engagement of volunteers that played the three proposed MLGs.

The next step is now the opening of the authoring tool to the staff of the botanical part. This step will be eased by having complete MLG scenarios corresponding to their need available as working examples.

References

1. Lee, H., Parsons, D., Kwon, G., Kim, J., Petrova, K., Jeong, E., Ryu, H.: Cooperation begins: encouraging critical thinking skills through cooperative reciprocity using a mobile learning game. Comput. Educ. **97**, 97–115 (2016)
2. Sharples, M., Arnedillo-Sánchez, I., Milrad, M., Vavoula, G.: Mobile learning. In: Balacheff, N., Ludvigsen, S., de Jong, T., Lazonder, A., Barnes, S. (eds.) Technology-Enhanced Learning, pp. 233–249. Springer, Dordrecht (2009). https://doi.org/10.1007/978-1-4020-9827-7_14
3. Rosyid, H.A., Palmerlee, M., Chen, K.: Deploying learning materials to game content for serious education game development: a case study. CoRR abs/1608.01611 (2016)
4. Wandersee, J.H., Schussler, E.E.: Preventing plant blindness. Am. Biol. Teach. **61**(2), 82–86 (1999)
5. Bertrand, S., Cerutti, G., Tougne, L.: Bark recognition to improve leaf-based classification in didactic tree species identification. In: VISAPP 2017–12th International Conference on Computer Vision Theory and Applications, February 2017
6. Marfisi-Schottman, I., Gicquel, P.Y., George, S.: Meta serious game: supporting creativity sessions for mobile serious games. In: Proceedings of the European Conference on Game Based Learning ECGBL, Paisley, United Kingdom, pp. 407–415, October 2016
7. Marfisi-Schottman, I., Gicquel, P.Y., Karoui, A., George, S.: From idea to reality: extensive and executable modeling language for mobile learning games. In: Proceedings of the European Conference on Technology Enhanced Learning, Lyon, France, pp. 428–433, September 2016
8. Steinberg, D., Budinsky, F., Paternostro, M., Merks, E.: EMF: Eclipse Modeling Framework. Eclipse Series, 2nd edn. Addison-Wesley, Upper Saddle River (2009)
9. Torrente, J., Moreno-Ger, P., Fernández-Manjón, B., Sierra, J.L.: Instructor-oriented authoring tools for educational videogames. In: ICALT 2008, pp. 516–518. IEEE (2008)
10. Nielsen, J.: Usability Engineering. Morgan Kaufmann Publishers Inc., San Francisco (1993)
11. Karoui, A., Marfisi-Schottman, I., George, S.: Mobile learning game authoring tools: assessment, synthesis and proposals. In: Bottino, R., Jeuring, J., Veltkamp, R.C. (eds.) GALA 2016. LNCS, vol. 10056, pp. 281–291. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-50182-6_25
12. Holden, C.: ARIS: augmented reality for interactive storytelling. Mobile Media Learning, pp. 68–83. ETC Press, Pittsburgh (2015)
13. Perry, B.: Gamifying French language learning: a case study examining a quest-based, augmented reality mobile learning-tool. Procedia-Soc. Behav. Sci. **174**, 2308–2315 (2015)
14. Zbick, J., Nake, I., Jansen, M., Milrad, M.: MLearn4web: a web-based framework to design and deploy cross-platform mobile applications. In: Proceedings of the 13th International Conference on Mobile and Ubiquitous Multimedia, pp. 252–255. ACM (2014)

Investigating the Design and Evaluation of Educational Games Under the Perspective of Player Experience

Alysson Diniz dos Santos[1,2(✉)], Francesco Strada[1], and Andrea Bottino[1]

[1] Dipartimento di Automatica e Informatica, Politecnico di Torino, Torino, Italy
{francesco.strada,andrea.bottino}@polito.it
[2] Instituto Universidade Virtual, Universidade Federal do Ceara, Av. Humberto Monte, s/n, bloco 901, Fortaleza 60440-554, Brazil
alysson@virtual.ufc.br

Abstract. Educational games are praised for their potential in providing engaging and self motivating learning experiences. However, in order to fully leverage those capabilities, game designers have to carefully balance factors that provide an enjoyable experience, without losing the focus on the instructional aspect. This work examines the design and evaluation of educational games focusing on the perspective of individual player emotions and experience. This investigation targets the following research questions: how to include (and balance) both the educational and the player experience elements into the game design? Which specific methodologies are available to evaluate player experience in educational games? After discussing possible answers to these questions, we briefly sketch directions that demand further research.

Keywords: Educational games · Player experience
Serious game design · Serious game evaluation

1 Introduction

When focused on learning, serious games must provide an experience in which entertainment and instruction are seamlessly integrated. This requires to take into consideration different aspects during the design process. On one hand, well-designed games need to be grounded on learning theories and instructional strategies. These allow researchers to manipulate key variables and determine which factors have the greatest effect on learner motivation and achievement.

On the other hand, there seems to be a general consensus in the literature that, in order to effectively convey educational contents to players, educational games (EGs) need primarily to be fun, entertaining and engaging [25,31]. In other words, EGs should immerse players, delivering an experience in which their attention is fully turned to the desired content, which, in turn, creates the ideal situation for learning to happen [14]. However, how to reach this objective

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 218–227, 2017.
https://doi.org/10.1007/978-3-319-71940-5_20

is a long debated question among researchers and practitioners, which involves multiple interrelated aspects (such as appropriate feedback, user interface, narrative, challenge design and flow [32]) and disciplines (such as computer science, design, pedagogy and cognitive psychology).

For decades now, researchers attempted to offer adequate guidelines on how to produce "enjoyable" EGs [20]. Initially, these studies were focused on technological aspects, such as game interface design, interaction devices and usability issues [22]. Recently, there has been a shift towards providing a broader view over the player-game relation. Thus, going beyond the aforementioned technological aspects, current studies take into account as well several elements related to the personal and individual gaming experience. These factors can be clustered in different levels [32]: *behavioral* (how player behave during interaction, *e.g.* whether they laugh, smile or frown), *physiological* (how the game affects the player physiological activity, *e.g.* measuring variation of heart rate and blood pressure during the game), and *psychological* (related to the subjective individual experience, the actions driven by intrinsic motivation, the feeling of presence and immersion, and so on). Such a broader view over the game interaction is usually referred to as *player experience*.

The objective of this work is to discuss EGs design under the lenses of PX. We specifically address the following research questions: how to include (and balance) both educational and PX elements into the game design? Which specific methodologies are available to evaluate PX in EGs? We believe this discussion can be of interest for both researchers and practitioners, since, on one side, the adoption of mechanisms eliciting PX might facilitate the achievement of the instructional goals [32] and, on the other hand, it might suggest how to tackle a well–known issue, *i.e.* the lack (in many theoretical frameworks) of practical guidelines for harnessing immersion in EGs [7].

In the literature, several studies (most of which are listed in [32]) related PX and (general) entertainment games. However, these works do not take into account the learning component, which is fundamental for EGs. To this end, the recent work of Nagalingam and Ibrahim [24] analyzed several theoretical frameworks specifically addressing the relationship between PX and EGs. We think that the major contribution of our work with respect to [24] is the attention on the assessment of PX in EGs, which is fundamental to highlight the validity of the proposed design frameworks.

In the following Sections, we first discuss the concept of player experience. Then we detail the theoretical frameworks that correlate PX and EGs, and we discuss their evaluation under the PX perspective. Finally, we draw the conclusions.

2 Player Experience Definition

In order to structure a grounding terminology for this study, it is necessary to (i) (briefly) clarify the use in this work of the term player experience, since there seems to be some ambiguity in the literature [32], and (ii) describe its relations with other similar terms, like usability, playability, and game experience.

In human-computer interaction (HCI) field, *usability* is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [18]. This definition led authors to derive mathematical bases for usability factors, *i.e.* measuring effectiveness in terms of completion rates and errors, efficiency from time on task, and satisfaction through standardized questionnaires [27].

In a similar way, researchers adopted the term *playability* (or *game usability*) as a proxy for usability in games. In other words, playability refers only to tangible elements at the technological level (such as the game interface, the input and output devices and the design choices affecting the gameplay), which are usually evaluated (again) through objective quantitative measurements [22].

However, there are elements missing in this picture. A game is made to be *experienced* by the player and the thinking, feeling and effect on the individual need to be taken into account as well. Therefore, PX builds upon playability to encompass the domain of experiences made by the player while interacting with the game [12]. While some authors adopt the term *game experience* to denote this concept, we prefer to use PX since it highlights the central role of the player as the driver of the game design and evaluation.

PX has some similarities with *user experience* (UX), which studies how a person perceives and responds to the interaction with a system [5]. In our view, the main difference is that PX places higher attention in addressing the "emotion" dimension, which is central in describing and enhancing the interactive experience human enjoy at play (see also [17]).

3 PX Theoretical Design Frameworks for EGs

The interplay between PX and game–based learning is characterized by an intricate collection of relationships among several factors, such as playability, player context, cognitive and emotional states, instructional design and learning theories. All these factors need to be well–balanced within the game design. To this end, several works in the literature introduce models for guiding EG design under a PX perspective. In the following, we discuss those extracted from our literature review, whose protocol is detailed in the Appendix, with explicit reference to the relationship between PX and effective educational content delivery.

Game Experience Model (GEM). The core idea behind the GEM is that, to achieve the learning objective of the game, designers should take in account the social, temporal and spatial contexts [23]. This means that a game designer should work towards two joint objectives. First, refine and test the game software and balance the game variables in order to provide an optimal playability and, second, improve the overall PX by including specific elements, such as: the introduction of different player models (*e.g.* novice, experienced) and the use of adaptive game mechanics.

Empirical evaluation of GEM [23] showed its effectiveness in delivering an overall satisfactory PX, and prompting learning.

Flow Framework. Kiili *et al.* (2012) propose a framework to bring learners to the flow state, which is characterized by high focus, engagement, motivation, and immersion. Reaching this state guarantees that the player is completely engaged in the game activity, performing tasks effortlessly and distraction–free [2]. The framework defines a set of factors (the *antecedents*) needed to guarantee the flow emergence.

Clear goals aim at ensuring player's focus on the learning tasks. Thus, educational contents should be embedded in several goals (a main one presented at the beginning of the game, and smaller ones in the following). Another aspect influencing player's focus is *immediate feedback* of goal progression and achievement. Feedback pace must be optimal: too frequently will break the flow state, whereas too rarely will make information absorption inadequate.

Playability and *sense of control* are two interconnected factors. The first refers to the balance between the progress of challenge difficulty and player's developed skills. This aims at keeping the player in the flow state and "away" from both anxiety (challenge is too complex) and boredom (challenge is too easy). Such balance instills as well a sense of control to the player, which might enhance PX and absorption of educational context.

An assessment of the Flow framework has been attempted in [6]. In contrast with the hypothesis of [19], this work reported insignificant effect of flow with respect to the learning outcomes. However, since the author of [6] highlights some limitations of his work, these results are inconclusive and require further investigation.

Playability Model. Ibrahim *et al.* [16] detail a set of design guidelines to foster the *educational playability* of an EG (*i.e.*, a game being entertaining and educative at the same time). These guidelines include elements introduced by other frameworks, such as the concept of adaptation (GEM) and game goals, game control, and feedback (Flow). In addition, new concepts are introduced: (i) *ethics*: game contents should always be presented within an acceptable ethical framework; (ii) *realism*: the game should simulate real life scenarios, to facilitate content transfer between virtual and real worlds; (iii) *game reward*: the introduction of a reward system has positive effects on the player motivation and, as a consequence, on the learning outcomes; and (iv) *player knowledge*: players should be indirectly prompted to activate and use their prior knowledge, which should be enriched with new contents delivered throughout the game. A main drawback of this work is that neither [16] nor any work published afterwards detailed an actual use of the proposed design model.

3.1 Discussion

At first, we deem important to share an interesting similarity among the three analyzed frameworks: all of them suggest that EGs should provide individually tailored experiences (for instance, dynamically updating game difficulty). Authors claim that such dynamic adaptation may help players to achieve and maintain the flow state, a statement that seems to be confirmed by recent research results [26,29].

As another general comment, while all the described frameworks explicitly link PX and EGs, they all seem to be biased towards PX, and put lower emphasis on the instructional design. Indeed, all authors allege that a satisfactory PX is *ideal* for the learning to happen, but they do not further exploit *how* this learning should happen. Thus, several questions remain unanswered, such as: which educational contents are eligible to be taught, how they should be presented and which pedagogical approach is the most adequate for the intended educational objectives and target audience. On the other side of the coin, in the literature there are many design frameworks specifically aimed at guiding the instructional design of EGs (see [8] for an overview), which are biased towards the educational aspect, with little consideration for the PX.

In our opinion, both PX and instructional design are necessary conditions for providing effective learning in an interactive and engaging way. Therefore, we believe that EG designs should carefully balance both aspects. Actually, this can be done by adopting and "merging" two frameworks, one more focused on PX and another on instructional design. Clearly, our hope is the emergence of a model where these two aspect are balanced, a topic worth to be explored.

Another salient drawback is the lack of thorough validation of the proposed frameworks. Nevertheless, since all of them are based on sound theoretical research, we do not believe that their effectiveness is in question but, for sure, they would benefit from a solid scientific assessment.

It is also important to highlight that the described frameworks approach the learning domain from a broad perspective, without restrictions on game genres, platforms, and players' characteristics. However, learning is a complex and multi-faceted process, and players engage with particular games for a plethora of reasons (competition, collaboration, collecting prizes, stimulating creativity, and so on). Therefore, recent research focus on frameworks tailored to specific domains and target audiences. For instance, [4] presents a framework aimed at engaging young students (7 to 16 years) in mathematics EGs. Another example is in [15], where authors investigate the engagement of mobile games for teaching History to secondary students. Although promising, these approaches still require a thorough validation.

4 PX Evaluation in EGs

While sound theoretical design frameworks striving to enhance the integration between PX and instructional design are sorely needed, ensuring an effective and high–quality EG requires as well their joint evaluation. Such evaluation requires to combine several interrelated aspects: the game effectiveness in facilitating learning, its playability and the user engagement, fun and emotions. In the following, we discuss the tools that have been specifically proposed for the PX evaluation of EGs.

Heuristics. Heuristics are design guidelines that can serve as well as game evaluation tools. In evaluation mode, one or more double experts (*i.e.* usability specialists who ideally are also game players) analyze, according to the proposed guidelines, the features of the prototypical game and produce a list of usability problems that should be solved before the final game release.

Mohamed and Jaafar [21] detail the Playability Heuristic evaluation for Educational computer Game (PHEG), whose guidelines are grouped along five dimensions (interface, educational element, content, playability and multimedia). PHEG based evaluation should involve (at least) one specialist in each dimension, to offer multiple views over PX. In order to validate their model, authors cross referenced the problems identified by PHEG, with the ones highlighted by a panel of 115 University students while reviewing different EGs. In particular, the panel results indicated the effectiveness of PHEG in detecting flaws in the analyzed games [21], thus highlighting its efficacy as evaluation tool.

In another work, Fitchat and Jordaan introduce ten heuristics to evaluate PX in EGs, with focus on mobile devices. The guidelines refer to several aspects, such as: adequate use of technology, player's attention, learning content, in-game challenges and assistance during the interaction. To compile the guidelines, the researchers employed a qualitative method (audio recorded interviews) with participants playing StoryTimes, a game designed to teach multiplication to children [9]. The major drawback of the study is the lack of a sound evaluation of the proposed heuristics (StoryTimes was used exclusively to capture user feelings towards the interaction and not as the object of the evaluation).

Surveys. Pre and post-play surveys and interviews are probably the easiest and least expensive approach to PX assessment. Concerning them, we found in the literature only two questionnaires specifically conceived to analyze EGs: the EGameFlow and the Serious Game Experience Model.

EGameFlow [11] is a scale to measure learner's enjoyment in EGs that incorporates the learning dimension to the ones originally defined in GameFlow questionnaire (developed to analyze the flow concept in general video games [30]). The proposed scale contains 42 items spanning eight dimensions: concentration, goal clarity, feedback, challenge, control, immersion, social interaction and knowledge improvement. The data analysis demonstrated the statistical validity of EGameFlow (*i.e.* that the model is well–founded). However, since only four games of "low" complexity were analyzed, more research is needed to assess its general validity.

As highlighted in [3], one of the main drawbacks of the EGameFlow model is that it can only be applied to games with clearly defined learning outcomes. To tackle this issue, De Grove *et al.* proposed the *Serious Game Experience Model* (SGEM), which addresses games whose primary educational objective is more open/abstract. SGEM was tested in the evaluation of the game Poverty is not a Game (PING), whose aim is raising consciousness about poverty. SGEM is based on the Game Experience Questionnaire (GEQ), a tool used in the PX evaluation of several digital games [3]. Due to this, GEQ lacks of an evaluation of the educational part of the game, which was addressed in SGEM by introducing a Perceived Learning module. The PING game was evaluated by 340 third and fourth grade students (aged 14–16 years). Although the collected data indicated the statistical validity of SGEM, authors expressed their concerns about the model generalization, since it was tested with only one game.

Behavioural and physiological evaluation. Another validation approach is based on the analysis of data collected during the game [10]. In particular, two methods are emerging: using analytics and players' physiological measures.

Game learning analytics (GLA) refers to the integration of learning analytics approaches in EGs, *i.e.* collecting and analyzing data about players and their contexts, for purposes of understanding and optimizing learning [10]. According with Shoukry *et al.* [28], the adoption of GLA demands several steps. At first, it should be established the intended use of the extracted information, which can be used during game play (*e.g.*, to provide an adaptive and personalized learning experience or to help instructors directing the learning process) and/or after the game session (*e.g.*, to assess the learning outcomes). Then, it is needed to choose and capture data, *i.e.* define meaningful data to be captured and how to collect them. Finally, since data can be captured from various devices (*e.g.* in case of multiplayer settings) and can be of different nature (*e.g.* logs of players activities and video recordings), strategies to aggregate data should be defined.

Although the entertainment game industry has been collecting data from users for many years, EGs have not yet taken full advantage of it [28]. Therefore, even considering the growing available literature available, we could not find sound theoretical frameworks that have been thoroughly evaluated. For a more extensive discussion on this topic, the reader can refer to [10].

Physiological measures. Another objective technique is based on the measurement of physiological variables, which can be interpreted as indicators of player's cognitive and emotional states. In specific, for learning environments, attention, effort and excitement can be derived from posture, facial expressions, eye tracking, pupil diameter, skin conductance, heart rate, respiration and electro activity of derma and brain signals [28].

Several recent studies approached physiological assessment in the context of EGs. For instance, [13] found patterns in eye tracking data of students playing engineering design games (highest performing students looked at similar places for analogous amounts of time). In another example, [33] identified increased oxygenation and hemoglobin presence in certain areas of the cortex, while testers played learning games. In spite of the growing literature on the topic, there is still a need for frameworks able to merge the computational complexities of physiological approaches with theories for learning. A broader discussion on this topic can be found in [1].

4.1 Discussion

All the identified evaluation tools have pros and cons. For instance, heuristic evaluation is based on accepted principles and the empirical results obtained in various studies show its effectiveness in identifying major and minor game problems before its actual release. However, heuristics only foresee the mere involvement of expert users, thus neglecting to consider the end–users in the evaluation loop. As a consequence, this approach appears unable to fully capture and analyze player's emotions, which represent a relevant part of PX.

On the other hand, surveys are able to effectively capture subjective player preferences. Nevertheless, this method requires a sufficient number of respondents to be statistically significant, and may present discrepancies between objective and subjective user reactions (sometimes what players do is different from their claim on what they think they do). In addition, surveys heavily rely on player's memory and, therefore, information may be lost in the delay between action (gameplay) and recall (interview or questionnaire application).

In turn, behavioural and physiological evaluation allow data-driven analysis, which may help to clarify the complex mechanisms of learning while playing games. However, both approaches may increase assessment costs. Indeed, physiological assessment requires specific devices and controlled setups, and the (potentially) huge amount of recorded data might require the implementation of data mining approaches. In addition, data analysis in GLA might be troublesome, specially due to the (possible) heterogeneity of devices/sensors used to collect them. In a similar vein, physiological data are volatile, variable, and may be harsh to interpret.

Given their different characteristics, we believe that a more comprehensive assessment could be obtained by combining the analyzed methods, in order to simultaneously target playability (through heuristics), users' opinions (through surveys), behavioural and physiological data. While such multimodal approach could guarantee an effective PX assessment for EGs, it also creates barriers by increasing the complexity of the evaluation and, thus, a careful planning of such an experimental protocol should be taken into account.

Concluding, the analysis of the current state–of–the–art seems to reveal the need for thoroughly assessed evaluation tools capable of taking into account (at the same level) both the learning and PX dimensions. Thus, we think that tackling this issue in future studies would bring a relevant benefit for research in EG area.

5 Conclusion

The objective of this study was to investigate the design and evaluation of EGs under the PX perspective. As a first conclusion, the identified design frameworks for EGs lack a thorough evaluation and may possibly be extended, in order to offer a greater support to the instructional design. Finally, we briefly hinted that *data triangulation* (among heuristics, surveys and behavioral/physiological data collection) could be beneficial to achieve a thorough and joint assessment of player immersion and learning. The actual application of such evaluation to an EG, as well as the definition of a design framework balancing PX and instructional guidelines, would be interesting future extensions of this work.

Appendix: Literature Review Protocol

In order to draw a representative picture of the state of the art, we formulated the following guiding research questions: (i) what recent papers approached PX in EGs?; (ii) what are the design frameworks expressly proposed to address PX in EGs? And (iii) which tools and methodologies are used to evaluate PX in EGs?

We carried out a search with the terms "player AND experience AND learning", in ACM Library and IEEE Explore portals, limiting the time span between 2006 and 2016. We then removed duplicated results and selected relevant papers according to the following criteria: (i) does the study consider player experience/user emotions? And (ii) does the study deals with EGs?

Finally, the remaining papers were read completely, and their references were also taken into account, thus expanding the list of analyzed works.

References

1. Blikstein, P., Worsley, M.: Multimodal learning analytics and education data mining: using computational technologies to measure complex learning tasks. J. Learn. Anal. **3**(2), 220–238 (2016)
2. Csikszentmihalyi, M.: Beyond Boredom and Anxiety. Jossey-Bass, San Francisco (2000)
3. De Grove, F., et al.: Towards a serious game experience model: validation, extension and adaptation of the GEQ for use in an educational context. In: Calvi, L., Nuijten, K., Bouwknegt, H. (eds.) Playability and Player Experience, vol. 10, pp. 47–61. Breda University of Applied Sciences, Breda (2010)
4. Dele-Ajayi, O., et al.: Learning mathematics through serious games: an engagement framework. In: Frontiers in Education Conference (FIE), pp. 1–5. IEEE (2016)
5. I. DIS. 9241-210: 2010. ergonomics of human system interaction-part 210: Human-centred design for interactive systems. International Standardization Organization (ISO), Switzerland (2009)
6. Dolar, D.P.: Experiments on flow and learning in games. Ph.D. thesis (2014)
7. Dos Santos, A.D., Strada, F., Martina, A., Bottino, A.: Designing collaborative games for children education on sustainable development. In: Poppe, R., Meyer, J.-J., Veltkamp, R., Dastani, M. (eds.) INTETAIN 2016 2016. LNICST, vol. 178, pp. 3–12. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-49616-0_1
8. dos Santos, A.D., Fraternali, P.: A comparison of methodological frameworks for digital learning game design. In: de De Gloria, A., Veltkamp, R. (eds.) GALA 2015. LNCS, vol. 9599, pp. 111–120. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40216-1_12
9. Fitchat, L., Jordaan, D.: Ten heuristics to evaluate the user experience of serious games. Int. J. Soc. Sci. Humanit. Stud. **8**(2), 209–225 (2016)
10. Freire, M., et al.: Game learning analytics: learning analytics for serious games. In: Spector, M.J., Lockee, B.B., Childress, M.D. (eds.) Learning, Design and Technology, pp. 1–29. Springer, Switzerland (2016). https://doi.org/10.1007/978-3-319-17727-4_21-1
11. Fu, F.-L., Su, R.-C., Yu, S.-C.: Egameflow: a scale to measure learners enjoyment of e-learning games. Comput. Educ. **52**(1), 101–112 (2009)
12. Gerling, K.M., et al.: Measuring the impact of game controllers on player experience in FPS games. In: 15th International Academic MindTrek Conference, pp. 83–86. ACM (2011)
13. Gomes, J., et al.: Analysing engineering expertise of high school students using eye tracking and multimodal learning analytics. In: Educational Data Mining (2013)

14. Gunter, G.A., Kenny, R.F., Vick, E.H.: Taking educational games seriously: using the retain model to design endogenous fantasy into standalone educational games. Educ. Tech. Res. Dev. **56**(5–6), 511–537 (2008)
15. Huizenga, J.: Digital game-based learning in secondary education. Ph.D. thesis (2017)
16. Ibrahim, A., et al.: Educational video game design based on educational playability: a comprehensive and integrated literature review. Int. J. Adv. Intell. Syst. **5**(3), 400–414 (2012)
17. Isbister, K., Schaffer, N.: Chapter 20 - the four fun keys. In: Game Usability, pp. 317–343. Morgan Kaufmann, Boston (2008)
18. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs), p. 45. The International Organization for Standardization (1998)
19. Kiili, K., de Freitas, S., Arnab, S., Lainema, T.: The design principles for flow experience in educational games. Procedia Comput. Sci. **15**, 78–91 (2012)
20. Malone, T.W.: What makes things fun to learn? heuristics for designing instructional computer games. In: 3rd ACM SIGSMALL, pp. 162–169. ACM (1980)
21. Mohamed, H., Jaafar, A.: Playability assessment for educational computer games: pilot study for model development. In: Zaman, H.B., Robinson, P., Olivier, P., Shih, T.K., Velastin, S. (eds.) IVIC 2013. LNCS, vol. 8237, pp. 396–407. Springer, Cham (2013). https://doi.org/10.1007/978-3-319-02958-0_36
22. Nacke, L.: From playability to a hierarchical game usability model. In: Proceedings of the Conference on Future Play, pp. 11–12. ACM (2009)
23. Nacke, L., et al.: Methods for evaluating gameplay experience in a serious gaming context. Int. J. Comput. Sci. Sport **9**(2), 1–12 (2010)
24. Nagalingam, V., Ibrahim, R.: User experience of educational games: a review of the elements. Procedia Comput. Sci. **72**, 423–433 (2015)
25. Prensky, M.: Digital Game-Based Learning. McGraw Hill, New York (2001)
26. Sampayo-Vargas, S., et al.: The effectiveness of adaptive difficulty adjustments on students' motivation and learning in an educational computer game. Comput. Educ. **69**, 452–462 (2013)
27. Sauro, J., Kindlund, E.: A method to standardize usability metrics into a single score. In: Human Factors in Computing Systems, pp. 401–409. ACM (2005)
28. Shoukry, L., Göbel, S., Steinmetz, R.: Learning analytics and serious games: trends and considerations. In: Proceedings of the 2014 ACM International Workshop on Serious Games, pp. 21–26. ACM (2014)
29. Shute, V., Ke, F., Wang, L.: Assessment and adaptation in games. In: Wouters, P., van Oostendorp, H. (eds.) Instructional Techniques to Facilitate Learning and Motivation of Serious Games. AGL, pp. 59–78. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-39298-1_4
30. Sweetser, P., Wyeth, P.: Gameflow: a model for evaluating player enjoyment in games. Comput. Entertain. (CIE) **3**(3), 3–3 (2005)
31. Whitton, N.: Learning with Digital Games: A Practical Guide to Engaging Students in Higher Education. Routledge, New York (2009)
32. Wiemeyer, J., Nacke, L., Moser, C., 'Floyd' Mueller, F.: Player experience. In: Dörner, R., Göbel, S., Effelsberg, W., Wiemeyer, J. (eds.) Serious Games, pp. 243–271. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40612-1_9
33. Witte, M., et al.: Neuronal correlates of cognitive control during gaming revealed by near-infrared spectroscopy. PLoS ONE **10**(8), 1–15 (2015)

The Effect of Uncertainty and Quality Perception on the Usage of Forecasting Tools – A Game Based Analysis

Richard Lackes, Markus Siepermann(✉), and Georg Vetter

Technische Universität Dortmund, Otto-Hahn-Str. 12,
44227 Dortmund, Germany
{richard.lackes,markus.siepermann,
georg.vetter}@tu-dortmund.de

Abstract. Forecasting tools are used in many areas of application to improve the decision making. But often, decision makers do not follow the forecast. This behavior is only desirable in situations when the forecast is faulty or does not consider all information. Otherwise, it bears suboptimal decisions. This paper investigates under which circumstances decision makers rely on or discard forecasts. In particular, the degree of uncertainty is in the focus. Results show that the more uncertain the decision situation is, the more a decision maker relies on the forecast if it contradicts his own estimation. The more the situation seems to be clear, the more he relies on his own estimation. Then, a faulty forecast can even confirm the decision maker to opt for a wrong decision.

Keywords: Decision making · Forecast · Forecast quality · Uncertainty

1 Introduction

Nowadays, software tools like expert systems or simulations are used in most planning processes to support the decision making process and to improve the final outcome [16, 26]. The term planning refers to decisions and actions to be taken in order to adapt to the future environmental development [3]. As future developments are not deterministic, forecasts are needed to predict and judge future outcomes depending on the given environmental conditions and possible actions [15]. Based on such projections into the future, the best decision alternative can be chosen. That means that in general two parties are involved in the planning process: The forecasting tool and the decision maker who uses the result of the forecast in his planning. Therefore, the quality of a decision depends on the one side on the quality of the forecast and on the other side on the discernment of the decision maker and his use of the forecast. In particular the correct use of the forecast results is highly dependent on the decision maker [2]. Even if a forecast is as accurate as possible, a decision maker may deviate from the forecast's advice be it due to additional information that was not considered in the forecast, be it due to subjective beliefs or personal convictions of the decision maker [4, 5, 24, 28].

But overriding a forecast is desirable only in a few specific situations, for example if the forecast is faulty, does not take all information into account, or is outdated.

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 228–237, 2017.
https://doi.org/10.1007/978-3-319-71940-5_21

Apart from that, deviating from a forecast's advice is not in accordance to the concept of the Homo oeconomicus. Therefore, it is crucial to know when and why decision makers do not use a forecast but rely on their own personal assessment of a situation instead. In general, forecasts are found to have an impact on the behavior of decision makers
[6, 9, 21, 27]. But the role of its quality could not be confirmed exactly by now. Some papers found that the quality of a forecast significantly influences how the forecast is used by decision makers [25], some papers could not confirm these findings [19]. However, the reasons why forecasts are used or not are investigated scarcely [1]. Therefore, this paper aims to shed light on the following research questions:

RQ: In which situations do decision makers rely on or discard forecasts?

Thereby, we focus on the uncertainty of the situation on the one side and on the quality of the forecast on the other side. Usually, decision situations differ concerning their degree of uncertainty. Some situations are quite unequivocal and only a small rest of uncertainty remains, some are much more difficult to assess. Depending on the situation, the quality of a forecast can therefore be assessed easily or not. To study these situations, we use the simple game High or Low. There, the decision maker is exposed to a simple decision situation. Depending on a game card made visible to him he has to choose if the next card that will be revealed is higher or lower than the card he can see. In addition, a forecast is presented to the decision maker that - depending on the likelihood - gives an advice if the next card is higher or lower. To study the influence of a forecast's quality, the forecast is manipulated and gives a wrong advice in some situations. This setting allows us to study the influence of forecast quality and of the degree of uncertainty on the usage of the forecast.

The remainder of this paper is therefore organized as follows: Sect. 2 gives an overview about the theoretical background and the related literature in this field. In Sect. 3, we explain the methodology of our study and in particular the game High or Low. We analyze the behavior of the participants in Sect. 4 and discuss the results in Sect. 5.

2 Theoretical Background

There are different streams of literature that analyze forecasts and the use of forecasts by decision makers. The first one deals with the interpretation of forecasts that is often not done correctly. According to Juanchich and Sirota [10] more than 50% of people participating in their survey did not interpret a forecast correctly. Maines and Hand [14] showed that individuals are not able to do a proper forecast because they cannot weigh time series information correctly. Also the self-assessment of people is defective. In the study of Lucarelli et al. [13] participants should assess their risk tolerance level and perform different tasks. A high share of participants thought themselves to be risk averse but acted like a risk taker.

The reason for this misalignment might be that the participants were not experienced with uncertain situations and the use of forecasts. Therefore, a second stream investigates the role of experience. These papers found a positive relationship between experience and the usage of forecasts [11, 12]. Experience reduces the emotional reactivity to events and therefore and improves the usage of forecasts during the decision making process.

The third stream of forecast literature has a look at the forecast itself. The more a firm invests in the forecast, the higher the quality of the input data is, and the more efficiently the forecast is done, the better is the quality of the forecast's outcome [7]. This helps for the usage of the forecast. As Smith and Mentzer [25] have shown, the accuracy of a forecast has a positive effect on its usage. The more accurate a forecast is, the higher its quality is perceived and the higher the quality is perceived, the more the forecast is used. However, when a forecast is done manually by analysts, they pay more attention to single risks than to the general risk situation [22]. Therefore, several authors investigate to what extent different forecasts are perceived as useful and credible. If a forecast is in line with other key performance indicators, a second, independent forecast is perceived as less useful. But if there is a misalignment between forecast and prior trends, decision makers mistrust this forecast [21] and prefer a second, independent forecast [6]. The different effects of forecasts can be employed by firms. If forecasts of future outcomes are reduced, this dampens the expectations of investors so that possible losses in the future are reduced [9].

However, as the effect of a forecast highly depends on the decision situation, a third stream investigates the characteristics of the decision situation. If decision makers face a quite comfortable situation, they tend to disregard forecasts even if the forecast is perceived as reliable and accurate [19]. Only if the situation is risky or if the decision makers expect a risky situation in the near future, they make use of the forecast. Unfortunately, O'Connor et al. [19] could only explain about 20% of the variance. This indicates that there are more factors to be examined.

Therefore, this paper also investigates the decision situation and its influence on the usage of forecasts. It is most related to the one of O'Connor et al. [19]. But in contrast to them, we focus on short-term decision marking instead of long-term decisions. In addition, we also vary the quality of the forecast. To some extent, this setting equals the one of Gaynor and Kelton [6]. The forecast provided in our study can be compared to the one of the firm and the personal assessment of the decision maker to the one of the analyst.

Besides the behavioral research stream, the paper also contributes to gamification in empirical studies as it uses game elements in a serious non-game context to motivate player's behavior [17, 23]. Nevertheless, the number of studies applying gamification to empirical studies, as a special form of a non-game context, is scarce. Beside Musthag et al. [18] who used points and a leader board to compare players in their study and papers in game theory using the Prisoner's Dilemma Game to investigate risk strategies and behavior (e.g. [8]), this is the first study in which a simple card game was used to analyze human behavior.

3 Methodology

To answer the question how decisions are influenced by the uncertainty of the situation, we used a simple card game called high or low. The game was expanded by a forecast which was shown to the participants. The forecast was either correct calculated or manipulated. Hereby, the participants could be confronted with selected situations in an observed laboratory experiment. These situations are characterized by their degree of uncertainty. Before we describe the forecast in more detail, the following section will give a deeper understanding of the game design.

3.1 The High or Low Game

The card game high or low can be played with a 32 or 52 card deck. To reduce complexity of the analysis we choose a 32 cards deck. The game is played in rounds. In each round two cards are drawn from the pile of unused cards. One card is shown to the player while the other card remains hidden. The player has to guess if the hidden card will be higher or lower as the shown card. All cards can be differentiated by their color and number/face. The sequence of cards from high to low is firstly dependent on the number or face. The order is from high to low: ace, king, queen, jack, 10, 9, 8, 7. The order of colors are from high to low: clubs, spades, hearts, diamonds.

Before a player makes a decision, a forecast is shown. The forecast as a decision support system can make three recommendations: higher, lower or uncertain. The recommendation relies on the calculated probability of the chances that the next card will be higher or lower. If the forecast tells "uncertain" the chances are too close to make a reliable recommendation. The calculation will be described in Sect. 3.2.

The player can choose if the next card will be higher or lower. Therefore, all players are forced to give an answer, even if the forecast shows "uncertain". After the player made a decision the second card is shown. If the player made the right decision, one point was added to his account. If the decision was false, no points were removed. The player can see his own point account and his overall high score. On a leaderboard, all players were able to compare their high scores with each other.

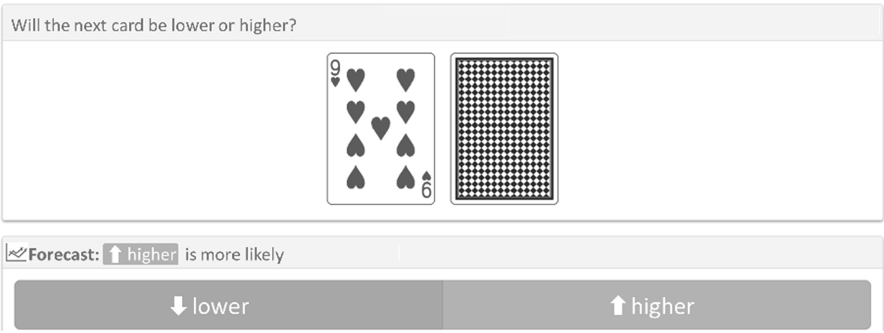


Fig. 1. Screenshot of the game

At the end of each round both cards were put to the pile of used cards. Since the deck contains 32 cards, a deck lasts 16 rounds. After 16 rounds the pile of used cards was shuffled and taken as the pile of unused cards for the next 16 rounds. The players were aware of the decreasing number of cards in the pile of unused cards and when the pile of used cards was shuffled. The game was implemented as an online game in a web browser. Figure 1 shows a screenshot of the game.

3.2 The Forecast

As mentioned above a decision support system was implemented in the form of a forecast to help the player. The forecast calculates the probability that the next card will be higher. This probability depends on the number of unused cards with a higher value as the shown card and the overall number of cards remaining in the pile of unused cards. Therefore, the forecast remembers all played cards. The following examples will help to explain the forecast. Example A: If the player gets the 9 of hearts as first card in the deck, 22 cards have a higher value and 31 cards are remaining in the unused pile. The probability that the next will be higher is 22/31 $\approx$ 0.71. Example B: During three rounds only cards below 9 of hearts were drawn. Then, when 9 of hearts is drawn, the probability will be 22/25 = 0.88. In both examples, the forecast would be "higher". Since the forecast is not deterministic but stochastic, the forecast does not have to be right. Hence, we implemented the prediction "uncertain" which will reduce the negative perception of the forecast. In situations where the probability is close to 50% a false prediction is more likely. By predicting "uncertain" no false prediction was given, but even no right prediction was given either.

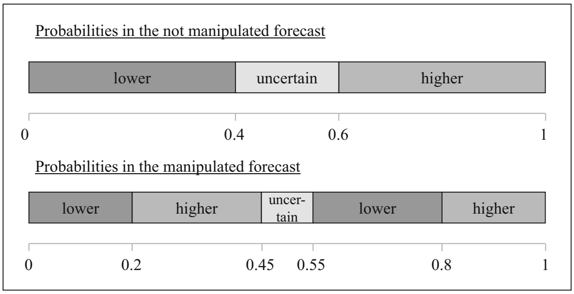


Fig. 2. Probabilities in the forecast

During the first two decks, the forecast predicts according to the calculated probabilities to build trust in the forecast. Afterwards the forecast is manipulated so that we could examine the behavior of the players confronted with a manipulated forecast in situations with different degrees of uncertainty. The manipulated forecast is also based on the probabilities used for the correct calculated forecast. We manipulated the forecast by switching the prediction from "higher" to "lower" and vice versa. There can be cases where it would be obvious that the forecast is manipulated. For this reason, we did not manipulate predictions with a probability above 80% and below 20%. According to the examples above: In example A the prediction would have changed from "higher" to "lower" and in example B the prediction remains as "higher". If a participant stopped playing and returned later, he started again with the correct calculated forecast for the first two decks. To produce more situations with a manipulated forecast we shortened the range of the uncertain prediction. Figure 2 shows how the manipulation was implemented.

4 Analysis

The laboratory experiment took place in the beginning of 2016. Mainly students participated in the experiment who are said to be good surrogates of managers [20]. In total, 10,202 data records were collected from 112 participants. Each data record represents a playing round and contains the drawn card for the player, the hidden card, the calculated probabilities, and the number of played decks. 2,626 of these data records had a manipulated forecast and could therefore be used in the analysis.

As we aim to examine the influence of the uncertainty on the behavior of players to follow a faulty forecast, we need a proxy for how uncertain the player judges the situation regardless of the forecast provided. As such, we use the naïve forecast that only takes into account the currently drawn card and does not consider the cards drawn previously during the deck. That means the naïve forecast sets the number of cards above the drawn card card in relation to the number of all cards of the deck, regardless of used cards. Thus, the probability of the naïve forecast in example A and B is equal ($22/31 \approx 0.71$). Then, two cases can be distinguished: (a) the faulty forecast equals the naïve forecast and (b) the faulty and the naïve forecast are distinct. For the analysis, two key figures are calculated: the percentage of players following the forecast in a situation of uncertainty and the degree of uncertainty of the situation.

The percentage of following players is calculated for every possible situation of uncertainty and each case by dividing the number of situations in which the manipulated forecast was followed by the number of times the situations appeared. The uncertainty of the situation is dependent on the calculated probabilities mentioned above. The closer the probabilities are to 50%, the more uncertain the situation is. Therefore, we map the difference between the calculated probabilities and 50% to the interval [0,1] where 1 represents a complete uncertain situation (probability of 50%) and 0 a complete certain situation (calculated probability is 0% or 100%).

For each case, a linear regression model is applied with the percentage of situations where the manipulated forecast was followed as the dependent variable Y and the degree of uncertainty as independent variable X. By identifying the unknown parameter β_1, it is possible to say how the degree of uncertainty influences the percentage of followed manipulated forecasts.

$$Y = \beta_0 + \beta_1 X \tag{1}$$

First, we have a look at case a) where the provided forecast equals the naïve forecast that we use as a proxy for the players' estimation of the uncertainty of the situation. Figure 3 shows the regression.

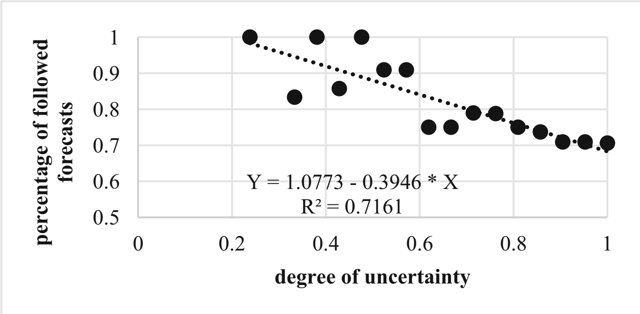


Fig. 3. Percentage of followed forecasts when provided and naïve forecast are equal

The manipulated forecast reassures the estimation of the players. As we can see when both predictions concur, the percentage of followed manipulated forecasts is high. The percentage decreases with an increasing degree of uncertainty. The slope of the function is quite small ($\beta_1 = -0.3946$) and the coefficient of determination is quite high ($R^2 = 0.7161$). There is one situation where the player chooses even with low degree of uncertainty not to follow the manipulated forecast Y A possible explanation for this data point can be that the player disagreed with the manipulated forecast since he did not use the naïve forecast for his own prediction. According to the correct calculated prediction, the behavior of the player in this situation is reasonable. Since this is the only outlier, we ignore this phenomenon.

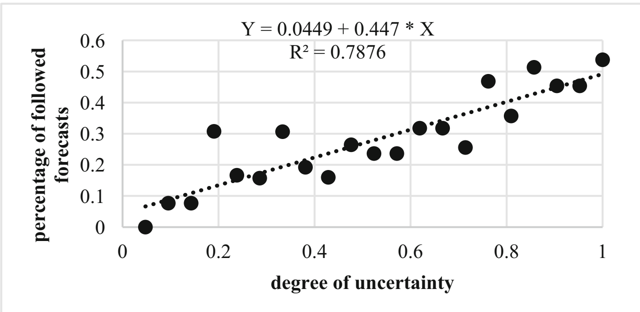


Fig. 4. Percentage of followed forecasts when provided and naïve forecast differ

When we focus on the situations where the estimations of the players are not in line with the manipulated forecast, we get the results depicted in Fig. 4. In this case, the player perceives a conflict between his own estimation and the forecast of the game. The percentage of followers is smaller than in the first case. The degree of uncertainty has a positive effect on the percentage of followed forecasts. Here, the slope is a bit higher than the slope in the first case ($\beta_1 = 0.447$) and the coefficient of determination as well ($R^2 = 0.7876$). The results confirm our assumption that the decisions of the players mostly rely on the naïve forecast.

5 Conclusion

5.1 Discussion

In this study, we aimed to identify the influence of the degree of uncertainty on the decision to follow a (manipulated) forecast. For this, we divided the analysis into situations where the player's prediction is in line with the forecast or not. It was shown that the degree of uncertainty impacts the percentage of followed forecasts. The different slopes of the functions show the diverse effect of the degree of uncertainty in both cases. If the forecast of the game does not agree with the player's estimation, the percentage of followed forecasts exceeds the 50% mark only in very uncertain situations. In line with [21], the player has an overall distrust towards the forecast until the uncertainty is too big.

When the manipulated forecast equals the estimation of the player, his prediction gets reassured and he perceives no conflict between both predictions. Therefore, we could confirm Maines and Hand [14] that the players were unable to make a good estimation based on prior situations. Here, the percentage of followed manipulated forecasts never drops below 70%. Interestingly, with an increasing degree of uncertainty, players distrust the manipulated forecasts. A possible explanation could be that some players calculated the correct probabilities in uncertain situations and based their decision more on this key figure. Nevertheless, most of the players followed the manipulated forecast even when the situation was very uncertain.

Both results show that players can distinguish between certain and uncertain situations and that they get unalienated when the degree of uncertainty increases. In uncertain situations, players tend to consider a forecast that differs from their own estimation. But it is also possible that the players did not consider the manipulated forecast in clear situations at all, as found in [19].

5.2 Implications

Expert and decision support systems are often a fundamental part of the decision making process in firms. The study can help to understand the drivers of the decision makers who follow a false prediction. It should increase the awareness for situations where decision support systems can have disadvantages. Also, decision makers can reflect on their own situation in terms of uncertainty to avoid wrong decisions. Decision support systems should be used with care since their performance can differ through the time as in our study, especially in a dynamic and stochastic environment which almost every firm is confronted with.

To reduce the percentage of followed manipulated predictions, additional information can be given to the decision maker, e.g. the probabilities of the possible predictions. The transparency of the expert systems and decision support systems can also be increased to reduce the number of wrong decisions.

For research, this study has shown that simple games are suitable for the observation of human behavior in general and decision makers in particular. The game has motivated many participants to take part and remain in the experiment.

5.3 Limitations

As usually, the study has some limitations. The number of observed players was quite small since 70 players did not generate enough data records to be in the analysis. All participants are German. To eliminate a possible cultural bias the experiment should be repeated with international participants. Further, we assumed that all players use a naïve prediction for their own decision making process. Even if the decision of the player and the forecast of the game are the same, the player does not have to have followed the prediction. Thus, in some cases the conformity can be coincidence. At last, the wrong decisions had no effect since no points were subtracted from the high score. False decisions in a business environment can have drastic consequences.

5.4 Future Work

In future work, we want to expand the experiment to gather more data records. By this we not solely want to validate present results but to get more insights about the decision making process. First, an analysis of different types of decision makers should be done with respect to demographic data. Also, the criteria of the decision making process of the participants should be investigated. In upcoming experiments, we will eliminate the prediction “uncertain” to force the decision maker to agree or disagree with the forecast.

References

1. Aziz, R.A., Manap, J.A.: Forecasting practice and perception of forecasting in a Government-Funded Malaysian University. J. Acc. Bus. Manage. **2**, 94–107 (2008)
2. Berinato, S.: What went wrong at Cisco. CIO Mag. **8**, 52–62 (2001)
3. Deshpande, R., Zaltman, G.: Factors affecting the use of market research information: a path analysis. J. Mark. Res. **19**, 14–31 (1982)
4. Fildes, R., Hastings, R.: The organization and improvement of market forecasting. J. Oper. Res. Soc. **45**(1), 1–16 (1994)
5. Fishbein, M., Ajzen, I.: Predicting and changing behavior: the reasoned action approach. Taylor & Francis, New York (2011)
6. Gaynor, L.M., Kelton, A.S.: The effects of analyst forecasts and earnings trends on perceptions of management forecast credibility. Acc. Finance **54**(1), 189–210 (2014)
7. Glaum, M., Schnürer, K., Schmidt, P.: What determines managers’ perceptions of cash flow forecasting quality? Evidence from a multinational corporation. J. Int. Financ. Manage. Acc. **3**, 298–346 (2016)
8. Hogan, J.L., Risher, R.H., Morrison, B.J.: Social feedback and cooperative game behavior. Psychol. Rep. **34**(3), 1075–1082 (1974)
9. Huang, W.: The use of management forecasts to dampen analysts’ expectations by Chinese listed firms. Int. Rev. Financ. Anal. **45**, 263–272 (2016)
10. Juanchich, M., Sirota, M.: How to improve people’s interpretation of probabilities of precipitation. J. Risk Res. **3**, 388–404 (2016)
11. Lo, A.W., Repin, D.V.: The psychophysiology of real-time financial risk processing. J. Cogn. Neurosci. **14**(3), 323–339 (2002)

12. Lo, A.W., Repin, D.V., Steenbarger, B.N.: Fear and greed in financial markets: a clinical study of day-traders, No. w11243. National Bureau of Economic Research (2005)
13. Lucarelli, C., Brighetti, G., Uberti, P.: Misclassifications in financial risk tolerance. J. Risk Res. **4**, 467–482 (2015)
14. Maines, L., Hand, J.R.M.: Individuals' perceptions and misperceptions of time series properties of quarterly earnings. Acc. Rev. **71**, 317–336 (1996)
15. Mentzer, J.T., Bienstock, C.C., Kahn, K.B.: Benchmarking sales forecasting management. Bus. Horiz. **42**(3), 48–56 (1999)
16. Moon, M.A., Mentzer, J.T., Smith, C.D.: Conducting a sales forecasting audit. Int. J. Forecast. **19**(1), 5–25 (2003)
17. Morford, Z.H., Witts, B.N., Killingsworth, K.J., Alavosius, M.P.: Gamification: the intersection between behavior analysis and game design technologies. Behav. Analyst **37**(1), 25–40 (2014)
18. Musthag, M., Raij, A., Ganesan, D., Kumar, S., Shiffman, S.: Exploring micro-incentive strategies for participant compensation in high-burden studies. In: Proceedings of the 13th International Conference on Ubiquitous Computing, pp. 435–444. ACM (2011)
19. O'Connor, R.E., Yarnal, B., Dow, K., Jocoy, C.L., Carbone, G.J.: Feeling at risk matters: water managers and the decision to use forecasts. Risk Anal. **25**(5), 1265–1275 (2005)
20. Remus, W.: Graduate students as surrogates for managers in experiments on business decision making. J. Bus. Res. **14**(1), 19–25 (1986)
21. Rupar, K.: Significance of forecast precision: the importance of investors' expectations. Contemp. Acc. Res. **2**, 849–870 (2017)
22. Sarens, G., D'Onza, G.: The perception of financial analysts on risk, risk management, and internal control disclosure: evidence from Belgium and Italy. Int. J. Discl. Gov. **2**, 118–138 (2017)
23. Seaborn, K., Fels, D.I.: Gamification in theory and action: a survey. Int. J. Hum Comput Stud. **74**, 14–31 (2015)
24. Simon, H.A.: Models of Man; Social and Rational. Wiley, New York (1957)
25. Smith, C.D., Mentzer, J.T.: User influence on the relationship between forecast accuracy, application and logistics performance. J. Bus. Logistics **1**, 159–177 (2010)
26. Stock, J.R., Lambert, D.M.: Strategic Logistics Management, vol. 4. McGraw-Hill/Irwin, Boston (2001)
27. Stone, D.N.: The joint effects of DSS feedback and users' expectations on decision processes and performance. J. Inform. Syst. **1**, 23–41 (1995)
28. Wheelwright, S.C., Clarke, D.G.: Corporate forecasting: promise and reality. Harvard Bus. Rev. **54**(6), 40–64 (1976)

Exploring Context-Aware Activities to Enhance the Learning Experience

Jannicke Baalsrud Hauge[1,2(✉)], Ioana Andreea Stefan[3], Antoniu Stefan[3], Massimiliano Cazzaniga[4], Pau Yanez[5], Tomasz Skupinski[6,7], and Francois Mohier[8]

[1] BIBA – Bremer Institut für Produktion und Logistik GmbH, Hochschulring 20, 28359 Bremen, Germany
baa@biba.uni-bremen.de
[2] Royal Institute of Technology, Mariekällgt. 3, 15181 Södertälje, Sweden
jmbh@kth.se
[3] Advanced Technology Systems, Str. Tineretului Nr 1, 130029 Targoviste, Romania
{ioana.stefan, antoniu.stefan}@ats.com.ro
[4] Imaginary, Piazza Caiazzo 3, 20124 Milan, Italy
massimiliano.cazzaniga@i-maginary.it
[5] Geomotion Games, Marie Curie, 8-14, 08042 Barcelona, Spain
pau.yanez@geomotiongames.com
[6] Ifinity, ul. Chmielna 2/31, 00-020 Warsaw, Poland
tskupinski@getifinity.com
[7] SGNS, ul. Towarowa 35/49, 00-869 Warsaw, Poland
tskupinski@sgns.pl
[8] ORT France, 24, rue Erlanger, 75016 Paris, France
francois.mohier@ort.asso.fr

Abstract. Mobile geolocation applications have been growing in popularity in the last decade. The ability to run a service on a mobile device that provides facts or recreational information to users opens up new opportunities to create engaging learning contexts. However, the potential of such services has not been fully exploited within educational settings, as compiling functional, student-friendly, context-aware learning journeys required advanced programming skills. The authors discuss this challenge and present tools that facilitate the construction of gamified learning paths, which integrate context-aware activities and minigames as motivational drivers for learning. The paper reports on the user testing feedback obtained in workshop settings.

Keywords: Authoring tools · Learning path · STEM · Minigames Beacons · QR codes · GPS · Learning experiences

1 Introduction

Technology-based innovation brings forth new opportunities to construct engaging and motivating learning journeys [1]. Context-aware services, in particular, provide means to expand the learning experiences outside of the classroom and produce learning

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 238–247, 2017.
https://doi.org/10.1007/978-3-319-71940-5_22

contexts that immerse students. By using beacons, a class of Bluetooth low energy devices that broadcast their identifier to smartphones, tablets and other devices to perform actions when in close proximity to them, teachers are able to model flexible and transformative learning sequences that provoke students to learn through exploration [2]. While digital games have been successfully used to engage students [3–6], context-aware learning activities provide a new approach, which takes advantage of context, place, and time of use, enabling teachers to increase the immersion level of learning.

Context-aware services represent a cost-effective solution that enables the implementation of agile and flexible teaching strategies that answer to the highly dynamic expectations of the new generations and support participation [7, 8]. However, it is necessary to empower the teachers with user-friendly tools that do not require advanced programming skills.

The paper addresses these challenges and presents a context-aware tool that interconnects with an Authoring Tool for Gamified Lesson Paths (AT-GLP) and a Metagame Runtime that are part of a pervasive learning platform. Together they enable the designing of innovative, game-based learning processes, which addresses the student expectations and have a great impact on the effectiveness of individual/group learning. These tools allow the designers to contextualize and build new context-based activity scenarios for implementation in a lesson path adapted to the different needs of its users (i.e. student and teachers). The paper outlines challenges and architecture of the Context-aware authoring system before it presents some of the minigames and finally details the feedback from the first usage in a workshop and lessons learned.

2 Authoring Context-Aware Activities

A set of tools has been developed in order to facilitate the construction of location-based learning systems that aim to engage students through new experiences that motivate learning. This section presents the architecture and functionalities of these tools. Learning Experiences (LEX) created using the Authoring Tool for Context-aware Challenges (ATCC) are reusable and customizable and their purpose can be changed. Teachers and learning designers are able to reuse existing LEX or their own, just personalizing basic parameters. This gives them the possibility to change the number and places of the locations to visit, customize the difficulty level or change the pedagogical content to fit the context-based activity to new lesson plans in a quite easy way, without programming skills.

2.1 The ATCC Architecture

An Authoring Tool for Context-aware Challenges (ATCC) has been developed in order to enable teachers to create gamified learning paths. The ATCC consists of three main components:

- The configuration tool for the Context-aware Challenges
- The Beacons inventory
- The Adobe Air Native Extensions (ANE) Extension

These components act as capability builders, enabling teachers to create playful location-based experiences. The ATCC uses three position tracking technologies: beacons, GPS, and Quick Response (QR) Code.

The ATCC communicates directly with the Authoring Tool for Gamified Lesson Path (AT-GLP). The ATCC serves as the main tool for context-aware learning activities and allows teachers to setup specific parameters of the experience. To complement ATCC, a Beacon Inventory was developed to store beacons and QR codes. It also helps teachers manage the position tracking technologies and assigning them to proper classes and groups. The Metagame Runtime is a customizable context-based activity environment that incorporates all the learning activities from a Gamified Lesson Path (GLP) in a game narrative. The Metagame Runtime is developed in AdobeAIR and requires the ANE native plugin in order to interact at low level with the device's functionalities transmitting position information from beacons to the Metagame application. Students access the Metagame to play the activities, including minigames, defined by the AT-GLP and the ATCC.

The diagram in Fig. 1 presents the high level architectural workflow of the communications between the ATCC, the AT-GLP and the Metagame.

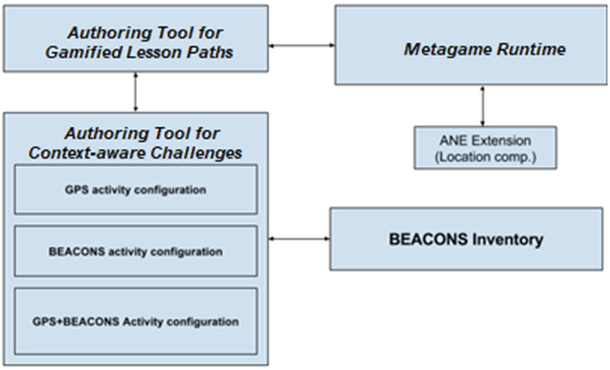


Fig. 1. The architecture of the ATCC

2.2 Creating Context-Aware Activities

After a GLP is defined in the AT-GLP, teachers can create context-aware games and activities in the ATCC. In the ATCC, teachers are also able to specify location for the learning activities in the GLP and define other parameters. The system uses a set of predefined activities to make activities, and no programming skills are required. This component together with the Beacons Inventory, allows the creation of mixed location based activities using both GPS and Beacons (or QR codes) (Figs. 2 and 3).

The context-aware activities can be part of any GLP and are created using an integrated location-based games authoring tool (Fig. 4).

It enables teachers to extend the students "game-based learning experience" beyond the boundaries of the classroom and allows them to explore the school, the surroundings of the school, and the neighborhood or the city, interacting with a real-life map.

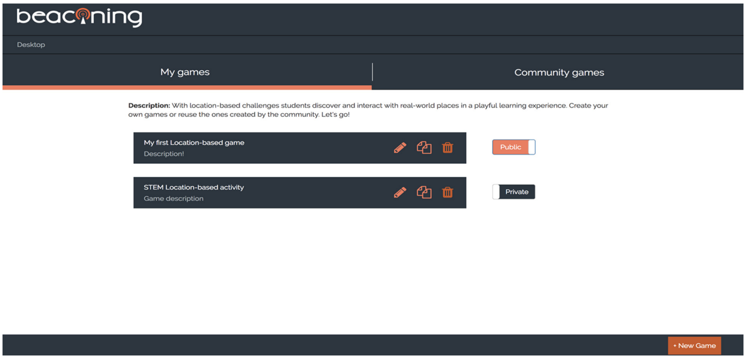


Fig. 2. The dashboard of the games used in the ATCC

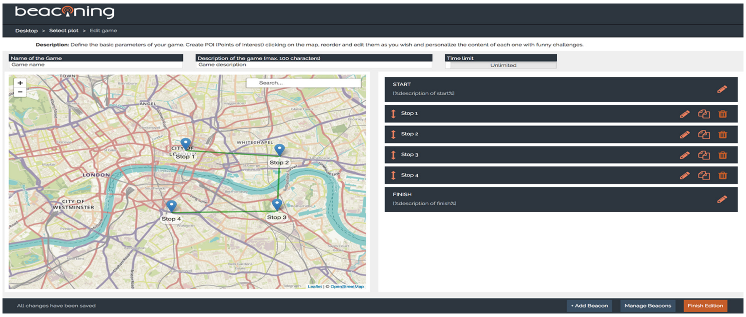


Fig. 3. Creating a "Follow the Path" context-aware activity

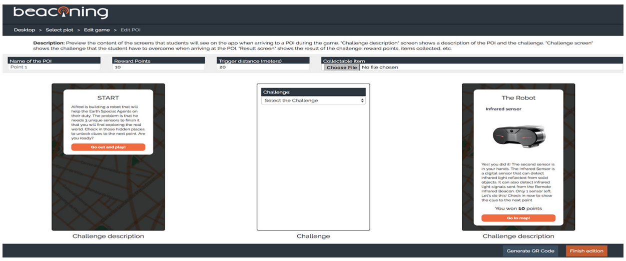


Fig. 4. Personalizing pedagogical content of the context-aware activity

Context-aware activities are a collection of POIs (Points Of Interests) where each point is associated with a learning activity like a minigame with a contextualized quiz, or a simple check-in option to verify the position of the user in the geolocalized mission.

2.3 Creation Pipeline

During the development of the ATCC, the aim was to keep the teachers experience as simple and intuitive as possible. Below there is a simplified diagram of the configuration pipeline of a context aware lesson path (Fig. 5):

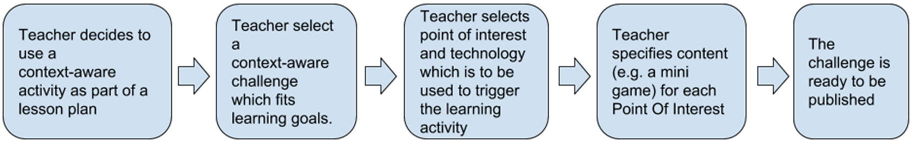


Fig. 5. Simplified creation pipeline graph

1. The teacher creating a lesson path decides if a context-aware activity fits with the pedagogical purposes of the educational activity;
2. The teacher selects a context aware activity option from the AT-GLP;
3. The AT-GLP opens the ATCC;
4. If there are any previously created or saved context aware activities, the teacher will be able to select an existing one and personalize it or create a new one;
5. When the teacher creates a new context aware activity, they can decide which technology it is going to use: GPS, BEACONS and QR Codes;
6. If the teacher selects the technology, the authoring tool will show the different context aware activity types;
7. The teacher decides which context aware activity fits the pedagogical purpose of the lesson path;
8. After selecting the context aware activity type the teacher can personalize for each step of the activity a minigame to be launched contextually to the current position of the student.
9. The tool enables to save work in progress that can be edited further before publishing an activity;
10. After finishing the configuration process, the activity can be published.

2.4 Examples of Context-Aware Activities

Context-aware services process information on the human factor (e.g. emotional state, habits, behaviors), and on the physical environment the user activates in, taking into account social interactions, co-location of others, group dynamics [10]. This context is explored by games to create compelling user experiences. Below there is a list of context-aware activities that can be used to gamify learning paths:

- *Follow the Path*: linear location-based game activity where students have to find and check-in at specific Points of Interest (POI). All points are shown in the map and the winner is the one who arrives first to the last POI.

- *Treasure hunt*: linear exploratory location-based game activity where the goal is to find a hidden treasure in the real world. Individually or in groups, students will have to find clues that give them information that has to be interpreted to find the final location of the treasure.
- *Capture the flag*: non-linear competitive location-based game where students are split into two different teams. Each team has a base where the flag is allocated. The goal of the game is to capture the enemy's flag and bring it to the base as many times as the game organizer decides.
- *Rat Race*: linear competitive location-based game where two or more teams of students have to participate on a race and be the first to reach the finish line. From the starting point to the goal the students have to solve activities in different POI's. The path of each team is different from the other.
- *Conquest*: non-linear competitive location-based game where students teamed up in different teams have to conquer different zones of the city while solving activities. Zone size and shape is automatically created by the context aware authoring tool. Zones can be reconquered by other teams. The first team to conquer a certain number of zones defined by the teacher wins the activity.
- *Jigsaw*: Linear competitive location-based game where the goal is to be the first team to arrive to a specific location overcoming activities in different POI. Different teams start the race at different POI and every time a team solves the activity of a POI a clue to the next one is shown (Fig. 6).

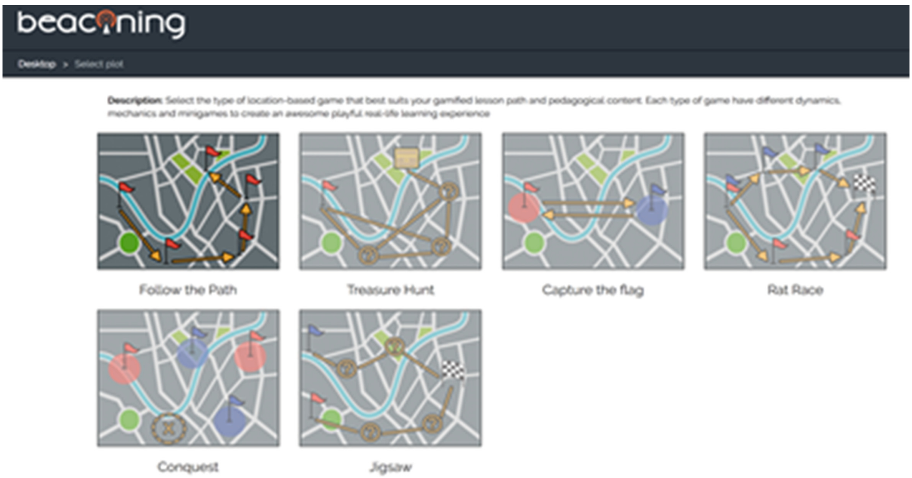


Fig. 6. Context-aware game types available on the Authoring Tool

3 Minigames

In a context-aware scenario, the learning path can integrate several minigames that are exposed to the learners as activities. When a learner reaches a certain location, a minigame is made available as a new activity. In game design, minigames are a special type of game activity, different from the main overall game experience that can address specific goals, according to choices made in the game's design. They can, for example, address specific topics or tasks engaging the player in a particular activity or can serve as a fun and entertaining experience inside the overall game story. In this second case we usually define them as "break games". From the user interface point of view, they typically provide a user interface with completely different graphic design from the main game play, highlighting that it is a different area within the game environment with a specific user interface depending on the specific aim and goal.

Mini games are short 2D web based game experiences that aim to engage the student with a different game play along with a fun activity. Minigames are inserted into a user experience workflow where a caller component (location based activity or overall meta-game) triggers the opening of a mini game, and at the end of the session, the player is sent back to the caller game. In the Fig. 7 below there are some screenshots of minigames with reusable and configurable game mechanics that have been developed to assess or train STEM skills.

RoboCode minigame

Swipe&Seek minigame

Fig. 7. STEM minigames

These are two of the games developed for BEACONING. The first screenshot is taken from the RoboCode minigame, a programming and robotics specific game mechanic where the player is asked to create an algorithm to get the robot from one end of the path to the other. The creation of the algorithm involves dragging and dropping the instructions from an instructions repository to an execution area, arranging them in the correct order. The second screen is taken from the Swipe&Seek minigame, where the player is asked to answer the question by finding and highlighting the answer in the grid by dragging a finger over the correct tiles with a "Ruzzle like" game approach.

To summarize, for all games the key features of the mini game components are:

- Reusability: to use the same game while addressing different STEM topics.
- Configurability: to allow teachers and learning designers to fully customize the mini game behaviors.

- Portability: due to web based architecture, the minigames can be part of a hybrid mobile architecture, running inside a native web view component, as well as part of a fully web based architecture, both on mobile and desktop browsers, due to the liquid and responsive layout of the user interfaces.

The next section reports on the first experience teachers and students have had with using the ATCC as a part of the lesson path as well as a context aware game.

4 First Piloting

This section demonstrates the first results of using different types of games in a workshop setting. The first workshop only reports on the usage of the context-aware game, whereas the second reports on the using different elements of the BEACONING system. More workshops will be carried out in autumn 2017.

4.1 First Experience of a Context-Aware Game

The context-aware games play an essential role in the lesson paths designed for creating higher engagement and pervasiveness among the users, so this part was tested first during the BEACONING platform Demo Session at the eLSE conference in Bucharest (April 2017). This scenario is related to the LEGO Mindstorm and the Coding & Robotic Lesson Path and used the Robot Treasure Hunt game (Sect. 2.2).

The location-based outdoor game using GPS technology was created applying "Treasure Hunting" game mechanics: a linear exploratory location-based game activity where the goal is to find a hidden treasure in the real world individually. Participants had to find clues on specific Points of Interest (POI) that gave them information that has to be interpreted to find the final location of the treasure. 12 students participated in the experiment. The gameplay was constructed as follows:

- Activity presentation: the game provided explanation of the game goal.
- Clue: the system sent users a clue with information about the next POI in real world. Users had to understand and move around to find the real location.
- Map: when users were trying to find the POI, the distance remaining was shown in real time to assist them with finding it.
- Activity screen: when users arrived to the right POI, the system automatically triggered a screen with the activity to overcome: minigames, check in, upload pictures, etc.
- Final screen: when users have found all the clues until the last POI, the system showed a final screen with the reward and information related to the storytelling.

Twenty-three students, aged 16 to 18, coming from different schools across Romania were divided into teams. After installing the app on their mobile phones (different types of devices), they went outside to discover the clues. The goal was to be the first group to return after solving all activities. The feedback was collected both using a questionnaire, as well as a part of an oral discussion. The questionnaire consisted of 20 questions that aimed to identify the main types of mobile devices used by students, playing habits (e.g. how often they play, what kinds of games they prefer), the "ideal game", and how students perceive pervasive learning. Students were enthusiastic about the context-aware, play-learning approach. They appreciated that context-aware games make learning fun, more engaging, and also easier. By integrating theory with practice, students have fun and learn without realizing it. Students considered that games help them develop their ability to retain more information. They suggested several subjects where the demoed game could be used (like for math (geometry), biology and history), and proposed improvements: the games should be more complex and more difficult to increase engagement; they should be more flexible to meet the different levels of knowledge.

4.2 Pilot Testing of Beaconing Components

Based upon the feedback from the students, the context-aware activity was improved to better fit the purpose in the lesson paths and used again. It was a part of a workshop with a pilot school aiming at engaging teachers to use the system and to show the opportunities created by designing more engaging and personalized lesson paths. On August 30 and 31, around 35 teachers from different ORT schools in France tried both minigames and a context aware activities designed according to the process described in Sect. 2.4. The location-based activity prepared for experimenting the new way of teaching used geolocalisation quests. 5 out of 35 teachers played the game in the same way as in Sect. 4.1, and then reported back to the other teachers.

The second part of the workshop was used to identify possible new Gamified Learning Paths, integrating the geolocalisation quests by imagining new ways to experience from their students. The feedback was collected in a debriefing section. Most of the teachers thought that Geolocation/Beacons provide new possibilities of improving the existing courses and extend the control of the teachers on the flow of the courses - inside/outside of the schools, thus they would be interested in elaborating new Learning Experiences further. However, they also made clear that they missed several functionalities and that we need to improve the usability, since it is not trivial to change neither the games nor the lessons paths.

5 Discussion and Future Work

This article describes the usage of an authoring system that enables teachers to integrate minigames and context-aware games as a part of lesson paths for teaching STEM subjects among students within the range of 16–24 years. It reports on the first experiences of around 50 teachers and students, where teachers have used the authoring system for adapting games to fit the individual lesson path and subject, and students have played a context-aware game. Even though the teachers reported that they would like to use it and the system is accepted as very innovative, it still requires quite much understanding of game design concept. In addition, they asked for a better usability with more functions. Based upon their feedback, we have added some functionalities that will be implemented and tested in November.

Acknowledgement. This research is partially funded under the Horizon 2020 Framework Program of the European Union, BEACONING – Grant Agreement 68676 and by Unitatea Executiva pentru Finantarea Invatamantului Superior, a Cercetarii, Dezvoltarii si Inovarii (UEFISCDI) in Romania, Contract no. 19/2014 (DESiG).

References

1. Huang, Y.M., Chiu, P.S.: The effectiveness of a meaningful learning-based evaluation model for context-aware mobile learning. Br. J. Edu. Technol. **46**(2), 437–447 (2015)
2. Hwang, G.-J.: Definition, framework and research issues of smart learning environments - a context-aware ubiquitous learning perspective. Smart Learn. Environ. **1**, 4 (2014). https://doi.org/10.1186/s40561-014-0004-5
3. Hamari, J., Shernoff, D.J., Rowe, E., Coller, B., Asbell-Clarke, J., Edwards, T.: Challenging games help students learn: an empirical study on engagement, flow and immersion in game-based learning. Comput. Hum. Behav. **54**, 170–179 (2016)
4. Wickham, C., Girvan, C., Tangney, B.: Constructionism and microworlds as part of a 21st century learning activity to impact student engagement and confidence in physics. In: Sipitakiat, A., Tutiyaphuengprasert, N. (eds.) Constructionism 2016, pp. 34–43. Suksapattana Foundation, Bangkok (2016)
5. Bodnar, C.A., Clark, R.M.: Exploring the impact game-based learning has on classroom environment and student engagement within an engineering product design class. In: Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality, pp. 191–196. ACM (2014)
6. Sabourin, J.L., Lester, J.C.: Affect and engagement in game-based learning environments. IEEE Trans. Affect. Comput. **5**(1), 45–56 (2014)
7. Petrolo, R., Loscri, V., Mitton, N.: Towards a smart city based on cloud of things, a survey on the smart city vision and paradigms. Trans. Emerg. Telecommun. Technol. **28**, 1–11 (2015). Wiley, http://dx.doi.org/10.1002/ett.2931, https://hal.inria.fr/hal-01116370
8. Dai, W., Liu, J.J., Korthaus, A.: Dynamic on-demand solution delivery based on a context-aware services management framework. Int. J. Grid Utility Comput. **5**(1), 33–49 (2014). 26
9. Curran, K.: Recent Advances in Ambient Intelligence and Context-Aware Computing, IGI Global (2014)
10. Gajjar, M.J.: Mobile Sensors and Context-Aware Computing, Morgan Kaufmann (2017). ISBN: 9780128017982 (electronic bk.), 0128017988 (electronic bk.), 9780128016602 (print)

Erratum to: Emotions Detection Through the Analysis of Physiological Information During Video Games Fruition

Marco Granato(✉), Davide Gadia, Dario Maggiorini, and Laura Anna Ripamonti

Department of Computer Science, University of Milan, via Comelico 39, 20135 Milan, Italy
{marco.granato,davide.gadia,dario.maggiorini, laura.ripamonti}@unimi.it

Erratum to: Chapter "Emotions Detection Through the Analysis of Physiological Information During Video Games Fruition" in: J. Dias et al. (Eds.): Games and Learning Alliance, LNCS 10653, https://doi.org/10.1007/978-3-319-71940-5_18

The presentation of Fig. 2 was incorrect in the original version of this chapter. The correct version is given in next page:

The updated online version of this chapter can be found at
https://doi.org/10.1007/978-3-319-71940-5_18

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. E1–E2, 2017.
https://doi.org/10.1007/978-3-319-71940-5_23

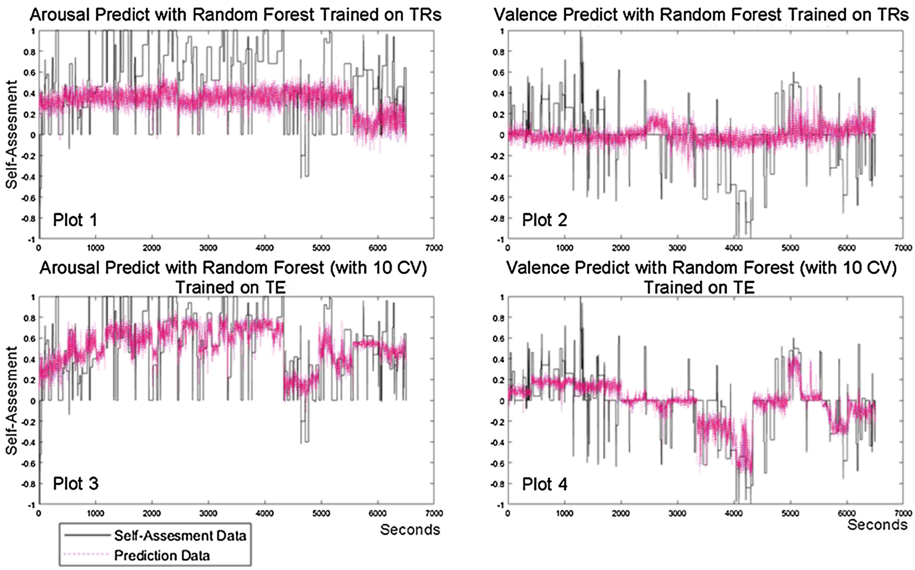


Fig. 2. The figure represents the mapped value of TE on emotional assessment. The time of each game session (using half second as step unit) is shown on x axis, the emotional assessment values are shown on y axis. The prediction values of the first two plots are defined by a Random Forest trained on TRs, while, the remaining two plots illustrate the predictions of a Random Forest trained with a 10 cross validation on TE.

The original chapter has been corrected.

Poster Abstracts

Using the Educational Potential Mapper to Design an Adaptive Serious Game: The “uManager” Case Study

Manuel Gentile(✉), Salvatore Perna, Giuseppe Città, Simona Ottaviano, Valentina Dal Grande, Dario La Guardia, and Mario Allegra

ITD - National Research Council of Italy, Via Ugo La Malfa 153, 90146 Palermo, Italy
{manuel.gentile,salvatore.perna,giuseppe.citta, simona.ottaviano,valentina.dalgrande,dario.laguardia, mario.allegra}@itd.cnr.it

Abstract. The frame pedagogy is a methodological framework that gives a key role to the concept of game mechanic in identifying the educational domain of a game. In this view, the activation of a subset of the game mechanics highlights a portion of the domain. Consequently, the game mechanics become a fundamental element in defining personalized and adaptive learning paths. As a case study, we discuss the implementation of a serious game called “uManager”, aimed at stimulating and supporting entrepreneurial skills in young people. In the paper, we provide a brief description of the process and the conceptual model behind the design and implementation of the game.

1 Introduction

In serious game design, the dynamic activation of Game Mechanics (GMs) is one of the main strategies to ensure the adaptability and personalized learning, since it allows the designer to modulate the game to adapt its complexity to the student’s knowledge.

A meaningful interpretation of the GMs dynamic activation strategy can be achieved applying the concept of the game mechanic presented in Gentile et al. [1]. According to the frame pedagogy approach [2], a GM is a verb in a context of use; consequently, it is possible to link each GM to a frame to obtain a network of connected frames that represent the domain knowledge of the game.

In this view, a partial activation of the GMs reduces the complexity of the game not only from a mere quantitative point of view but principally from a cognitive point of view. In fact, reducing the number of actions (GMs) on which the student has to reflect and decide allows students to focus on a portion of the domain knowledge elicited by the game. Therefore, a path of gradual activation of the game mechanics generates a sequence of activation maps which progressively tends to cover the entire game knowledge domain. Following this approach, we present the implementation of the uManager serious game, a management/construction game aimed at promoting entrepreneurial skills in young students (Fig. 1). The aim of the paper is to show how the

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 251–253, 2017.
https://doi.org/10.1007/978-3-319-71940-5

Fig. 1. A screenshot of the uManager game.

proposed methodology was used by the designer to create a game model that gives flexibility to the teacher in the process of learning design (personalization), setting the foundation for creating a fully adaptive game.

2 The Design Elements Behind uManager

uManager is a typical management/construction game aimed to foster entrepreneurial skills in young students in which the user has to construct and manage a holiday resort.

According to the framework presented in Gentile et al. [1], the adaptivity features of the uManager game was based on the *Game Mechanic* concept.

The game model connects the GMs with the *Levels* and *Tasks* to give to the game designer the maximum flexibility in the definition of a learning path. The *Level* concept is used to identify the set of enabled GMs. Each of activated GM is used to evaluate user activity according to a specific degree of difficulty. These evaluations are the base of the assessment of the resort quality. The *Task* concept is used to structure user activities; a task usually involves the use of one or more GMs and as such can be used to introduce new GMs to the player or consolidate the understanding of already known ones defining a quest. Finally, the *Game Model* defines the sequences of the levels and the progression mechanism, which can either be automatic or manually controlled by the teacher. In the first case, there are two main strategies to regulate the level progression; the former is to use rules linked to the evaluation of the user's performances, while the latter is achieved connecting the level progression to the virtual time progression.

References

1. Gentile, M., Città, G., Ottaviano, S., La Guardia, D., Dal Grande, V., Allegra, M., Jarvinen, A.: A semantic frame approach to support serious game design. In: Games and Learning Alliance, pp. 246–256. Springer International Publishing (2016)
2. Perna, S., Lo Monaco, F., Città, G., Dal Grande, V., Gentile, M., La Guardia, D., Ottaviano, S., Allegra, M.: Educational potential mapper - a semantic-based tool for analysing and designing serious games. In: Proceedings of 11th International Technology, Education and Development Conference, pp. 8298–8307 (2017)

Serious Games for Participatory Design, Crowdsourcing and Remote Usability Testing

Edward Oates(✉)

Cranfield University, Cranfield, UK
e.oates@cranfield.ac.uk

1 Introduction

Designers rely on direct access to 'users' to assist in their design process. Within defence Human Factors Integration, the engineering process of Human-Computer Interface (HCI) design sometimes has difficulty accessing users. This study investigates the extent to which Serious Games may offer an asynchronous remote alternative to 'face-to-face' design through Crowdsourced Participatory Design.

Usability testing of the HCI may also lack access to users. This is a serious limitation which may be offset by remote unmoderated usability testing. The extent to which Serious Games may support remote usability testing is also explored in this research.

This is a mid-study report on UK Ministry of Defence sponsored research under the Royal Navy 'DARE Innovation' initiative.

2 Aims

Assess the viability of a Serious Game to support a Crowdsourced Participatory Design process.

Assess the method of remote unmoderated testing with a Serious Game for selecting human-computer interface components based on Game-play scores.

3 Method

An Electronic Warfare interface was built within a maritime-air synthetic environment. The first-person web browser game emulated an existing naval helicopter HCI. The game task was to detect and identify ships using electronic support, radar and electro-optic controls and displays.

The game was hosted on the MOD intranet and made available for all 120,000 users to play. The security restrictions on write-back to the server prevented automatic data from being logged. There was no way of knowing how many played the game. Log Files were sent in attached to emails.

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 254–255, 2017.
https://doi.org/10.1007/978-3-319-71940-5

The Crowdsourcing task was to play the game, identify functions and displays that would assist game-play and recommend how these should be designed. All available methods of communication were used: phone, email, blog. It didn't matter what was requested as testing would later identify the useful and the not-so-useful. All ideas were built into the game no matter what the motivation; intrinsic or extrinsic.

Remote unmoderated testing was implemented through competitions with Amazon tokens as prizes. For the Level 3 Game, seven experienced operators competed using the Original Game and Crowdsourced Game interfaces. The data log files recorded were then analysed for correlations between player selections and game results.

An on-line questionnaire was used for System Usability Scale (SUS) assessment, Flow (Csikszentmihalyi) assessment and subjective assessment of how realistic the game was. The questionnaire results have yet to be analysed.

4 Results

The Crowdsourcing activity with the Serious Game engaged with a geographically dispersed population. Design ideas were freely offered to develop the functions and displays. Twenty-two developments were specified and added over 28 months.

Remote unmoderated testing took all function 'click' (selection) data, and display duration data and looked for correlation with Game-play results. Game-play results were scored on time, (shorter duration being better than longer duration) and points (a negative scoring system rewarded correct and penalised incorrect identification of shipping). Use of Inverse Synthetic Aperture Radar correlated with high-score/short-duration games. Some Crowdsourced functions may have assisted reduce durations.

5 Conclusions

Results show that Serious Games provide a credible tool for Crowdsourced Participatory Design with a geographically dispersed crowd of volunteers. Even with a minimally advertised game, the game generated actionable new design ideas.

Remote unmoderated testing using a Serious Game has provided some indication of which functions and displays underpin effective and efficient Game-play. Crowdsourced functions may have helped some lower score players to reduce their game-play times. Further data is required to provide greater certainty of this method.

Using Choreographies to Support the Gamification Process on the Development of an Application to Reduce Electricity Costs

Fernando Cassola[1,2](✉), José Iria[1,3], Hugo Paredes[1,2], Leonel Morgado[1,4], António Coelho[1,3], and Filipe Soares[1]

[1] INESC TEC - INESC Technology and Science, Porto, Portugal
{fernando.c.marques, jpiria, antonio.n.coelho, filipe.j.soares}@inesctec.pt
[2] UTAD – Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal
hparedes@utad.pt
[3] FEUP – Faculdade de Engenharia da Universidade do Porto, Porto, Portugal
[4] Universidade Aberta, Lisbon, Portugal
leonel.morgado@uab.pt

1 Extended Abstract

Building automation systems contribute to reduce electricity costs by managing distributed energy resources in an efficient way. However, a large share of consumption cannot be optimized through automation alone, since it mainly depends on human interactions. Gamification can be used as one form of changing users' behaviours [1], but its implementation does require assumptions on the behaviour patterns that need to be identified, encouraged, or discouraged. To tackle this problem, we propose a framework that joins building automation solutions with gamification techniques to enable behavioural demand response.

Ultimately various authors converge to define gamification as "the use of game design elements in non-game contexts" [2]. So we use game elements and mechanics to engage and motivate end-users on an interactive platform [3]. To persuade this approach we will follow the Six Steps Gamification framework [4]. Additionally employ another gamification design framework, which places more emphasis on human motivation: Octalysis [5], proposed by Yu-kay Chou.

Several authors [6] tell us that the "knowledge acquired in an action-based and meaningful context promotes behavioural change" [7]. So we propose identifying users' behaviours that can be potentially relevant in a three phase process. In the first phase we will be using the building automation systems to monitor electric consumption of all actions produced by the building equipment (elevators, air conditioning, etc.) and inhabitants. Next we intend to extract users' behavioural patterns as choreographies related with energy consumption. After that, analysing the energy consumption of choreographies, we will promote the most effective ones employing gamification techniques (or even promote novel choreographies), with the goal of achieving electricity savings.

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 256–258, 2017.
https://doi.org/10.1007/978-3-319-71940-5

Employing platform-independent choreographies is a way to guarantee the interoperability and integration with other systems and approaches. First of all, because we have multiple input datasets from which to create the choreographies of the building devices and their occupants. Namely we are going to use the data given by the building automation electricity meters and also use smartphone and desktop applications to collect indoor users' locations and behaviours. Platform independence will enable the future association/adaptation of choreographies to different inputs. The other reason is because we must be able to render them in various output systems. We need to employ a graphical interface, showing the collected behaviours associated energy consumption data, so human users can identify choreographies on that historical data registry. Choreographies will also be used on the gamified application in order to encourage users to adopt specific behaviours. So the platform independence enables choreographies to be treated as core data for this software solution, rather than a mere visual gimmick. For this purpose, we will follow the Ontology-based transformation approach for choreographies proposed by Silva et al. [8].

Concluding, this is an approach of using choreographies to support the gamification design process on the development of an application to reduce electricity costs. The primary obstacle to this field is to identify main behaviours and join them with the electricity meters. Another technical challenge is to draw alternative behaviours to reduce electricity costs. Rather the technical barriers we need to obtain a large spectrum of related behaviours and engage the building occupants to participate on this project. Using the gamification techniques presented merged with the independent choreographies method we can expect to present a new way to engage people to reduce energy costs.

Acknowledgments. This work is financed by the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project SmartGP/0003/2015 under the framework of ERA-Net Smart Grids Plus.

References

1. Burke, B.: Gamify: how gamification motivates people to do extraordinary things. J. Chem. Inf. Model. **53**(9), 1689–1699 (2013)
2. Deterding, S., Sicart, M., Nacke, L., Hara, K.O., Dixon, D.: Gamification. using game-design elements in non-gaming contexts. In: Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI EA 2011 (2011)
3. Seaborn, K., Fels, D.I.: Gamification in theory and action: a survey. Int. J. Hum. Comput. Stud. (2015)
4. Mathiassen, L., et al.: For the win: how game thinking can revolutionize your business. J. Syst. Softw. **38**(2), 398–424 (2012)
5. Chou, Y.: Octalysis – complete gamification framework. Website (2015). https://goo.gl/rjb1yq

6. Cordero, E.C., Todd, A.M., Abellera, D.: Climate change education and the ecological footprint, Bull. Am. Meteorol. Soc. (2008)
7. Lee, J.J., Ceyhan, P., Jordan-Cooley, W., Sung, W.: GREENIFY: a real-world action game for climate change education, Simul. Gaming (2013)
8. Silva, E., Silva, N., Morgado, L.: Staging choreographies for team training in multiple virtual worlds based on ontologies and alignments. In: Shumaker, R., Lackey, S. (eds.) VAMR 2014. LNCS, vol. 8526, pp. 105–115. Springer, Cham (2014)

Gaming Against Violence: An Exploratory Evaluation Through Mechanical Turk of the Efficacy of Persuasive Digital Games in Improving Unhealthy Relationship Attitudes

Drew Crecente[1(✉)] and Ruud S. Jacobs[1,2]

[1] Jennifer Ann's Group, Atlanta, GA, USA
[2] Erasmus Research Centre for Media, Communication, and Culture, Erasmus University Rotterdam, Rotterdam, The Netherlands

1 Introduction

Teen dating violence (TDV) is a pervasive and complex public health problem. Despite its ubiquity – an estimated 40% of graduating college students in the US have been in an abusive relationship at least once [1] – and adverse effects on, among others, academic achievement and substance abuse, 16 million (of about 21 million) US students are educated in states lacking teen dating violence (TDV) education legislation [2]. This causes institutions that lack in-house expertise on this matter to acquire TDV prevention programs that do not adequately prevent violence perpetration or address bystander intervention [3]. It is therefore necessary to approach this issue in a way that does not rely on institutional education practices. Digital games are a novel method to confront the problem of TDV and since 2008 Jennifer Ann's Group (JAG) has produced videogames for this purpose through its "Gaming Against Violence" program [4]. This technology solution allows school systems to affordably and reliably meet student needs without specialized training, knowledge, or additional resources. It also allows adolescents seeking to assist themselves or their peers to utilize a medium that promotes private exploration of sensitive topics at the player's pace. The current study explores the influences teen dating violence games have on players and their effectiveness at overcoming those challenges which limit traditional interventions.

Approaching dating violence prevention through a game allows creators to present their message in a wholly unique way. Although video games are still often seen as mere entertainment, their ability to effect both generally positive changes in players [5] as well as more specific changes in attitudes towards social issues [6] is increasingly recognized. By offering simulated environments, TDV games promotes experiential learning in players [7, 8] allowing them to reach their own conclusions about the critical aspects of TDV [9]. They also learn about this topic in the context of biographical or fictional stories, rather than in an abstract way [10]. Teens prefer games over other media, [8, 11] and see the use of computers in instruction as a more positive experience than other methods [12]. Though the games affect knowledge and skills, they also aim to change players' attitudes, to convince players certain behaviors constitute abuse and should not be condoned. This makes TDV games *persuasive games*.

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 259–262, 2017.
https://doi.org/10.1007/978-3-319-71940-5

While persuasive games have been shown to positively affect player-held beliefs on several topics, [6, 13, 14] evidence is still needed to support attitude-changing effects of TDV games [15]. Our study was therefore guided by the following research question: Do teen dating violence games change the attitudes of unguided players towards this topic when compared to games unrelated to the topic?

2 Methods and Results

A 2 (TDV and entertainment control games) by 2 (pre- and post-test) mixed experimental design was applied to 86 US-based participants (all adults, 76% over 25 years, 68% male) gathered through the Amazon M-Turk service [16]. To get an overview of the generalized efficacy of the TDV games, 5 different JAG games were included with each game taking 30–60 min to complete. Two browser-based games formed the control condition: Samorost, [17] and Today I Die [18]. This 'no-treatment' control served to eliminate re-testing and experimenter effects to show if the TDV games changed attitudes. The Attitudes Towards Dating Violence Scales [19] were used as the indicator of the games' effects.

In a repeated measures ANOVA, significant main effects were found for the pre- and post-test difference ($F(1,84) = 7.1, p = .009, \eta_p^2 = 0.8$), as well as for the condition ($F(1,84) = 6.7$, $p = .012$, $\eta_p^2 = 0.7$). An interaction effect was also found, indicating that TDV game players changed their attitudes towards the topic more than people who played the control game. The two conditions did not significantly differ at pre-test ($t(84) = 1.8$), though they were different at post-test ($t(84) = 3.3$, two-tailed-$p = 0.002$, Hedges' $g_s = .83$), with means indicating that the TDV players ($M = 4.6$, $SD = 0.5$) had greater scores than the control players ($M = 4.2$, $SD = 0.6$). When looking at the difference scores between the two observations, the increase was largest for players of the TDV games (Control $M = .2$, $SD = 0.3$, TDV $M = .0$, $SD = 0.2$, $t(84) = 2.5$, $p = 0.013$, Hedges' $g_s = .64$). The TDV games incited significantly greater attitude change in their players than the control games did. No significant differences were found among the 5 TDV games tested ($F(4,61) = .5$). Lastly, gender differences were found in the pre-test scores of players of the TDV games ($t(64) = 2.629$, $p = .011$, Hedges' $g_s = .67$), indicating women ($M = 4.7$, $SD = .4$) had higher scores than men ($M = 4.3$, $SD = .6$). The difference in effect between genders ($t(64) = 1.947$, two-tailed-$p = .056$, Hedges' $g_s = .50$) indicated that men's attitudes ($M = .2$, $SD = .3$) underwent greater changes than women's ($M = .1$, $SD = .3$), though this was not significant at .05 level.

3 Discussion

Summarizing, the interaction effect found indicates the change in attitudes for TDV game players was greater than for control game players. TDV game players changed their attitude positively, whereas the control group showed no increase. This means it is highly likely the TDV games had an effect on the attitudes of its players and games intended to have an effect on TDV attitudes affect these attitudes more than games not designed for this purpose.

To conclude, despite the fact that the TDV games used in this study were not developed with the current study's sample in mind, the games clearly affected their players. This result supports the viability of persuasive games in combating the all-too-common phenomenon of teen dating violence.

References

1. Forke, C.M., Myers, R.K., Catallozzi, M., Schwarz, D.F.: Relationship violence among female and male college undergraduate students. Arch. Pediatr. Adolesc. Med. **162**(7), 634–641 (2008)
2. U.S. Census Bureau. Source and Accuracy Statement for the October 2010 Current Population Survey Microdata File on School Enrollment, Washington (2011)
3. De La Rue, L., Polanin, J.R., Espelage, D.L., Pigott, T.D.: A meta-analysis of school-based interventions aimed to prevent or reduce violence in teen dating relationships. Rev. Educ. Res. **87**(1), 7–34 (2016)
4. Crecente, D.: Gaming against violence: a grassroots approach to teen dating violence. Games Health J. **3**(4), 198–201 (2014)
5. Granic, I., Lobel, A., Engels, R.C.M.E.: The benefits of playing video games. Am. Psychol. **69**(1) (2013)
6. Ruggiero, D.: The effect of playing a persuasive game on attitude and affective learning. Comput. Human Behav. **45**, 213–221 (2015)
7. Mellecker, R.R., Witherspoon, L., Watterson, T.: Active learning: educational experiences enhanced through technology-driven active game play. J. Educ. Res. **106**(5), 352–359 (2013)
8. Bowen, E., Walker, K., Mawer, M., Holdsworth, E., Sorbring, E., Helsing, B., Bolin, A., Held, P., Awouters, V., Jansz, S.: It's like you're actually playing as yourself': Development and preliminary evaluation of 'Green Acres High', a serious game-based primary intervention to combat adolescent dating violence. Psychosoc. Interv. **23**(1), 43–55 (2014)
9. Sung, H.-Y., Hwang, G.-J.: A collaborative game-based learning approach to improving students' learning performance in science courses. Com. Educ. **63**, 43–51 (2013)
10. Charles, D., McAlister, M.: Integrating ideas about invisible playgrounds from play theory into online educational digital games. In: Rauterberg, M. (ed.) ICEC 2004. LNCS, vol. 3166, pp. 598–601. Springer, Heidelberg (2004)
11. Blumberg, F.C., Blades, M., Oates, C.: Youth and new media: the appeal and educational ramifications of digital game play for children and adolescents. Z. Psychol. **221**(2), 67–71 (2013)
12. KovačEvić, I., Minović, M., Milovanović, M., de Pablos, P.O., StarčEvić, D.: Motivational aspects of different learning contexts: my mom won't let me play this game. Comput. Human Behav. **29**(2), 354–363 (2013)
13. Peng, W., Lee, M., Heeter, C.: The effects of a serious game on role-taking and willingness to help. J. Commun. **60**(4), 723–742 (2010)
14. Gerling, K.M., Mandryk, R.L., Birk, M.V., Miller, M., Orji, R.: The effects of embodied persuasive games on player attitudes toward people using wheelchairs. In: Proceedings, 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI 2014, pp. 3413–3422 (2014)

15. Jozkowski, K.N., Ekbia, H.R.: 'Campus Craft': A game for sexual assault prevention in universities. Games Heal. **4**(2), 95–106 (2015)
16. Amazon.com Inc. Amazon Mechanical Turk - Service Highlights. https://requester.mturk.com/tour
17. Amanita Design: Samorost. http://amanita-design.net/samorost-1
18. Benmergui, D.: Today I Die
19. Price, E.L., Byers, E.S., Belliveau, N., Bonner, R., Caron, B., Doiron, D., Greenough, J., Guerette-Breau, A., Hicks, L., Landry, A., Lavoie, B., Layden-Oreto, M., Legere, L., Lemieux, S., Lirette, M.-B., Maillet, G., McMullin, C., Moore, R.: The attitudes towards dating violence scales: development and initial validation. J. Fam. Violence **14**(4), 351–375 (1999)

A VR Cardiac Auscultation Examination Game

Mario Vargas-Orjuela[1], Alvaro Uribe-Quevedo[3](✉), David Rojas[2], Bill Kapralos[3], and Byron Perez-Gutierrez[1]

[1] Universidad Militar Nueva Granada, Bogota, Colombia
marvaror@gmail.com, byron.perez@unimilitar.edu.co
[2] The Wilson Centre, Toronto, ON, Canada
david.rojasgualdron@gmail.com
[3] University of Ontario Institute of Technology, Oshawa, ON, Canada
alvaro.j.uribe@ieee.org, bill.kapralos@uoit.ca

Abstract. Virtual simulation and serious games are becoming more prevalent in training by providing simulated, monitored and controlled scenarios. They provide medical trainees with the opportunity to acquire both cognitive and technical skills outside of the medical environment in an engaging, and cost-effective manner. Here we present the design and implementation of a mobile game for cardiac auscultation training Preliminary results indicate the mobile game is engaging and enjoyable.

Keywords: Cardiac auscultation · Serious games · Simulator

1 Introduction

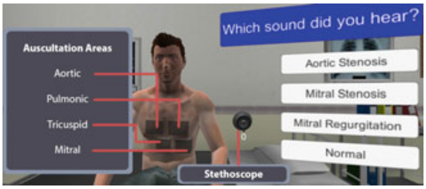
Cardiac auscultation is a routine medical examination procedure whereby a stethoscope hear and interpret heart murmurs [1]. Simulation, virtual simulation and serious games (that is, games developed specifically for education and training), and gaming technologies in particular, are capable of providing highly immersive, and engaging training experiences, not possible with traditional media. For example, Katz et al. (2017) developed a serious game to teach best practices for the anesthetic management of a standard orthotopic liver transplantation (OLT) procedure which itself is very complex [2].

In this paper we present a mobile-based serious game for cardiac auscultation examination, where a user diagnoses a virtual patient, and obtains scores for proper diagnoses, without penalties if mistaken. We compared the examination diagnosis between audio-only, a SAM 3G simulation manikin, and our game. We presented four random scenarios with different heart murmurs (i.e., normal, mitral stenosis, mitral regurgitation, and aortic stenosis), and virtual patients requiring a diagnosis from the user, by listening to the aortic, pulmonic, mitral and tricuspid chest areas. We chose the Unity3D engine to develop the serious game and integrate the game assets. We additionally incorporated learning and game mechanics including exploration, assessment, feedback, and repetition into the game [3]. We qualitatively measured engagement of our serious game using the verified Game Engagement Questionnaire [4].

J. Dias et al. (Eds.): GALA 2017, LNCS 10653, pp. 263–264, 2017.
https://doi.org/10.1007/978-3-319-71940-5

2 Results and Discussion

Seven fourth-year medical students with prior cardiac auscultation knowledge and experience and who considered themselves proficient with cardiac auscultation, volunteered to participate in the experiment. Participants completed a pre-test where none were able to identify the murmurs. During the testing, only two diagnosed the normal murmur with the simulation manikin, while with the game six students identified the normal murmur, four the mitral stenosis and the mitral regurgitation, and one the aortic stenosis (See Fig. 1). Based on the GEQ results, the serious game requires refinement to improve absorption and flow, while presence and immersion were perceived as engaging.



GEQ Element	Average Score
Presence	0.42
Absorption	-0.77
Flow	-0.04
Immersion	0.71

Fig. 1. In-game screen captures and GEQ results.

Future work will focus on additional game elements to increase flow and absorption. We will examine immersion and presence through the use of a virtual reality head-mounted display, and increase the number of participants to obtain statistically significant results.

Acknowledgment. The authors thank Universidad Militar Nueva Granada for funding Project ING INV-2363.

References

1. Chizner, M.A.: Cardiac auscultation: rediscovering the lost art. Curr. Probl. Cardiol. **33**(7), 326–408 (2008)
2. Katz, D., Zerillo, J., Kim, S., Hill, B., Wang, R., Goldberg, A., DeMaria, S.: Serious gaming for orthotopic liver transplant anesthesiology: a randomized control trial. Liver Transpl. **23**(4), 430–439 (2017)
3. Callaghan, M., Savin, M., McShane, N., Eguiluz, A.: Mapping learning and game mechanics for serious games analysis in engineering education. IEEE Trans. Emerg. Top, Comput (2015)
4. Brockmyer, J.H., Fox, C.M., Curtiss, K.A., McBroom, E., Burkhart, K.M., Pidruzny, J.N.: The development of the game engagement questionnaire: a measure of engagement in video game-playing. J. Exp. Soc. Psychol. **45**(4), 624–634 (2009)

Author Index