Sandro Hardy, Angelika Kern, Tim Dutz, Christoph Weber, Stefan Göbel, Ralf Steinmetz: **What Makes Games Challenging? - Considerations on How to Determine the "Challenge" Posed by an Exergame for Balance Training.**
In: ACM: Proceedings of the 2014 ACM International Workshop on Serious Games , ACM, November 2014.

What Makes Games Challenging? - Considerations on How to Determine the "Challenge" Posed by an Exergame for Balance Training

Sandro Hardy
Technische Universität Darmstadt
Multimedia Communications Lab (KOM)
64283 Darmstadt, Germany
sandro.hardy@kom.tu-darmstadt.de

Angelika Kern
Technische Universität Darmstadt
Institut für Psychologie
64283 Darmstadt, Germany
kernangelika@gmail.com

Tim Dutz
Technische Universität Darmstadt
Multimedia Communications Lab (KOM)
64283 Darmstadt, Germany
tim.dutz@kom.tu-darmstadt.de

Christoph Weber
Technische Universität Darmstadt
Arbeitsgruppe für Arbeits- und Ingenieurspsychologie
64283 Darmstadt, Germany
weber@psychologie.tu-darmstadt.de

Stefan Göbel
Technische Universität Darmstadt
Multimedia Communications Lab (KOM)
64283 Darmstadt, Germany
stefan.goebel@kom.tu-darmstadt.de

Ralf Steinmetz
Technische Universität Darmstadt
Multimedia Communications Lab (KOM)
64283 Darmstadt, Germany
ralf.steinmetz@kom.tu-darmstadt.de

ABSTRACT

Exergames are expected to induce desired health-related effects by incorporating a players' bodily movements into their gameplay. The effectiveness of these games is determined by the compliance to the training goals and by the frequency and the duration of playing. The latter is influenced by the player's motivation and increasing it is therefore essential for the effectiveness of exergames. Findings suggest that one of the main elements influencing motivation through the game is the "perceived challenge", consisting of the difficulty level posed by the physical interaction required for playing the game on the one hand, and the difficulty posed by the game mechanic itself one the other. If challenge could somehow be measured and intentionally varied, it might be possible to achieve a specific "level of motivation" for a given user. In this paper we present an approach that aims to quantify the amount of challenge induced by the variation of individual features of an exergames. In addition, the psychological self-determination and goal-setting theories are brought into the context of game research and evaluation, suggesting that setting a personal goal might influence the outcomes, too. Therefore, the dependence of motivation, perceived difficulty and performance on specific level features as well as goal-setting are being tested, aiming to find whether the effectiveness of the training can be increased intentionally by changing level features or setting personal goals.

SeriousGames'14, November 7, 2014, Orlando, FL, USA.

http://dx.doi.org/10.1145/2656719.2656728.

Categories and Subject Descriptors

H.5.2 [**Information Systems**]: User Interfaces—*input devices and strategies*; J.3 [**Computer Applications**]: Life and Medical Sciences—*health*; K.8 [**Personal Computing**]: General—*Games*

General Terms

Measurement, Documentation, Human Factors

Keywords

Serious Games; Health; Exergames; Training; Personalization; Adaptation; Sensors; Motivation; Difficulty; Goal Setting; Self Determination Theory

1. INTRODUCTION

Lack of exercise is a world-wide and well-known problem causing more than half of the ten most severe diseases worldwide. Among them are cardiovascular diseases, diabetes, cancer and chronic respiratory diseases that cause two thirds of the 57 million deaths worldwide [12]. The risk of suffering from cardiovascular diseases is increased by a sedentary lifestyle and the consumption of high-calorie food. The latter can in part be compensated by a sufficient amount of exercising, which however often poses a problem for elderly people: Because of higher possibilities for inactivity and health impairments caused by strokes, heart-attacks, diabetes, et cetera, their mobility is decreased. This limits their ability to perform physical activity. As a result, many elderly people suffer from low muscular strength and lack of coordination, flexibility and balance, which also increases the risk of falls. Appropriate training options accepted by a broad range of users could counteract this phenomenon and increase the quality of life of the individuals and also lower the societal health costs at the same time. To achieve this goal, the game "BalanceFit" was developed to facilitate training for the elderly. In order to be able to maximize the motivation of the player induced by the game, it is essential to analyze how single elements of the gameplay influence motivation and perceived difficulty, and if those outcomes might also depend on other external concepts like goal-setting. Ideally, understanding the relations between single elements of the gameplay enables researchers to precisely control the motivation, difficulty and performance and through this, also the effectiveness of the training.

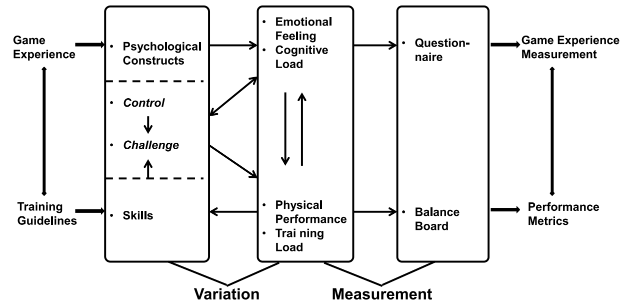

Figure 1: Relations between psychological constructs, requirements from sports science and their results.

2. RELATED WORK

The development of exergames involves research fields from at least three different domains: physical training, technology and game experience [6]. Sport science researches the effectiveness of exercises and training concepts and defines metrics for the evaluation of training effects, like the Tinetti-test [19] to measure gait stability and balance which are both indicators for the risk of falls. Multimedia technologies can deliver solutions to improve these concepts and physical assessments in a digital context, i.e., through measuring the center of pressure (abbr. COP), stability, range of movement, reaction time, or movement speed. Game experience research focuses on measuring the emotions of the users and developing concepts on how to influence these emotions.

Motivation is highly relevant if a certain task is to be completed (e.g. an exercise) because it concerns all aspects of activation and intention including energy, direction, persistence and equifinality [16]. When playing exergames, the player plays the game for training purposes (effectiveness) and because the game itself is motivating (attractiveness) [17]. Both attractiveness and effectiveness influence the frequency and duration of the training and therefore also the training load.

The theoretical model depicted in figure 1 sets the already identified constructs gathered from the scientific state of the art into the context of exergames. The motivational and physical aspects of exergames are denominated by game experience and training guidelines, which are considered as input parameters. They refer to psychological constructs and skills which result in emotional feelings, cognitive load, physical performance and training load. In both psychology and sport science, the term *challenge* is used and considered to be a highly relevant construct that affects motivation as well as the physical performance of players. The psychological and physiological effects can be measured by game experience measurements through questionnaires and performance metrics through sensors (such as a balance board).

There are many theories on what motivates users and different authors propose specialized concepts and design guidelines for motivation in games [18] and exergames. The concepts include - but are not the limited to - flow (challenge at the threshold between boredom and anxiety) [3], challenge, curiosity, feedback, control, empowerment and social interaction [6]. Other work shows the feasibility of digital virtual training and the configuration of exergames [11].

Perceived difficulty is one of the key elements for challenge, beyond other aspects such as feedback, social interaction, and competition [3]. In the case of exergames, perceived difficulty has to be divided into the difficulty of the physical interaction and the difficulty of the gameplay. Findings suggest that challenge influences motivation, and therefore the perceived difficulty is expected to affect motivation itself.

From a psychological point of view, two types of motivation can be distinguished: Behaviors pursued for their own sake are denominated "intrinsically motivated" and amongst other things lead to enhanced performance, persistence and self-esteem, whereas behaviors trying to achieve a certain positive end state or avoid a negative one are called "extrinsically motivated" [16].

The study of Vallerand and Losier [20] connects the psychological concepts of intrinsic and extrinsic motivation in sports-context. They refer to empirical data evaluating the cognitive, affective and behavioral effects of motivation and show a positive influence of motivation on performance, intensity and persistence.

A famous model to explain human motivation is the Self Determination Theory [16]. It suggests that there are psychological needs every human has which, if satisfied, lead to higher intrinsic motivation. These needs include the human need for competence (sense of efficacy, the perception that one is able to reach a goal), autonomy (volition and personal agency), and relatedness (social connectedness).

To be able to satisfy the need for competence, it is necessary to have the perception that one is able to reach a goal, thus a goal has to be set. The goal setting theory by Locke and Latham [9] states that conscious goals affect action by directing both cognitive/behavioral attention as well as effort towards goal-relevant activity. The authors cite empirical data which shows that the most difficult goals produce the highest levels of effort, performance and persistence, as long as there is a commitment to the task and as long as the limits of ability are not reached. Relevant for a persons' commitment is the perceived importance of the expected outcome and the belief to attain the goal. Therefore, if one wants to utilize the benefits of intrinsic motivation on a task (e.g. exergames), the stated possibilities of goals and needs should be kept in mind while designing or conceptualizing the games. The Self Determination Theory was implied on Video Games by Przybylski, Rigby & Ryan [13], stating that the players' experience of competence can be achieved by a balanced game difficulty and sufficient feedback of the actual performance. Autonomy can be influenced, among other things, through in-game choices over goals and strategies. There also have been some tendencies to include the psychological construct of goal-setting on persuasive technology [2, 10] to promote physical activities.

Based on these concepts, the following hypotheses were extracted to find out if one can influence motivation, perceived level difficulty or performance through level features or goal setting:

Hypothesis 1: Level features influence motivation.

Hypothesis 2: Level features influence the perceived difficulty.

Hypothesis 3: Level features influence performance.

Hypothesis 4: Goal setting influences motivation, perceived difficulty and performance.

3. METHODOLOGY

3.1 Evaluation Framework

The evaluation setting was created by using the previously developed evaluation framework StoryTecRT [7]. The framework allows to adjust certain game attributes while all other attributes keep the same characteristic, which allows to explain eventual differences in the results of those levels only by the characteristics varied. For measurements, the framework provides different loggers which record the raw input data of the used sensor (i.e. Balance Board, Ergometer, 3D-Tracker) as well as user-induced events (e.g. changing the sensitivity of the games sensorics), game-specific events (e.g. start and end of the game, reaching the goal, failing), game behavior (e.g. rotation angle, position of the ball) and physiological data (e.g. heart rate). The amount of trials or the time taken to complete a game can be taken as an indicator for the difficulty posed by a given level.

In this study, StoryTecRT was used to separate the game levels which fulfilled our needs for the operationalization of the selected constructs best. The relevant measures for this study were the raw sensor data from the four pressure sensors of the balance board, which allowed the user to interact with the game, and the game specific events, as start and end of the level defined the time needed to complete the level and the number of fails was used as an indicator for difficulty. The user was not allowed to change game parameters and no physiological data was logged.

3.2 Implemented Game

Based on the conceptual framework StoryTecRT, the exergame BalanceFit1[1] was developed, which focuses on the overall goal of fall prevention for elderly people. The game resembles classic wooden maze-and-ball-games and the goal of the game is to navigate a ball through a maze into a target. The maze is presented on a screen and the player steers by moving the the center of pressure of his body intentionally. The maze consists of walls which facilitate the task and game features which increase the difficulty of the game, such as round and rectangle open spaces to fall through or bridges and narrow passages which need higher accuracy in steering. As input control, a regular Nintendo Wii Balance Board is used. The Balance Board is a rectangle boards which has four pressure sensors, one at each of its edges, which allow to measure the weight distribution of the player standing on the board. From the weight distribution the COP is calculated which is used as input to control the game. If the player moves in a specific direction, the COP also moves in this direction. The virtual maze is rotated into the same direction and the virtual ball follows immediately. Because StoryTecRT allows the adaptation of the visual style, the layout of the labyrinth and the sensitivity of the controller, the system can be used by agile and fit users as well as by gait-impaired users (with a standing frame as hardware to provide stability) and even by wheelchair drivers [7].

[1]For details of the game please see www.spielend-fit.de

4. EVALUATION METHOD

4.1 Experimental Design

The evaluation was conducted through a two-factorial experimental design with repeated measurements. The first independent factor was goal-setting: The experimental group was asked to do a realistic but challenging estimation on the time needed to complete one level, whereas the control group was not asked to do such estimations. The second independent variable was the characteristic of the level parameters: 29 different features describing the different levels had been extracted from the game as possible candidates to influence the actions and reactions of a player. For this study, four of these features were selected. Due to the lack of a "ground truth", the selection is based on the assumptions of the authors, which are based on practical experiences from the usage of the system in a seniors' home during previous studies. The selected features, visualized in Figure 2, are:

- The number of bottlenecks (none or 1) being hindrances to be passed with at maximum 3 times the width of the ball (marked green in figure 2)
- The number of orientation changes necessary (ranged from 1-9) being the number of times a orientation change of 90° would be necessary to reach the goal (marked red in figure 2)
- The length of open curbs (none, short: 7.6 cm, medium: 14.5-17 cm) being the length of curbs where the ball is not secured by walls but could fall through the bottom of the board (marked orange in figure 2)
- The length of the shortest path (short: 15-17.5 cm, medium: 21.5-26.5 cm, long: 42.5-48.5cm) being the length of shortest path through the maze to reach the goal (marked blue in figure 2)

The levels which represent these features best were selected from the 52 existing levels of the game. The levels 5, 9, 10, 22, 45, 49, 50 were chosen (including level 7 as a training level), because in those levels the four said parameters were varied to an extent such that their influence on the perceived difficulty could be determined, while the other parameters were kept constant.

One of the dependent variables was the performance, operationalized by the time needed to complete the level. Another one was motivation, measured with a questionnaire for the acquisiton of the actual motivation [15] which evaluates motivation on the for scales *anxiety of failure*, *probability of success*, *interest* and *challenge*. The items aimed on the upcoming task (e.g. 'It is likely that I will not succeed.') were answered in a pre-questionnaire before the game; the ones reflecting on the conducted level in a post-questionnaire

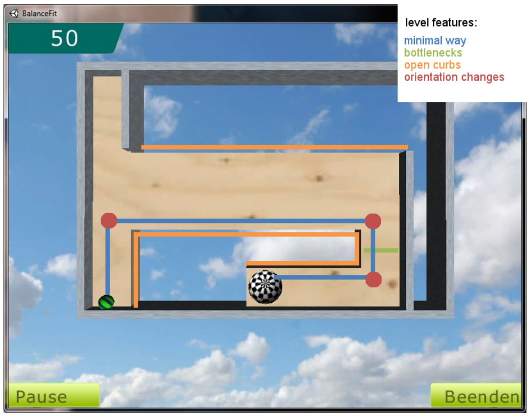

Figure 2: The chosen level features visualized in Level 50.

after the game. Furthermore, the subjectively perceived difficulty of the levels was measured as a dependent variable with the German translation [5] of the questionnaire 'subjective mental effort questionnaire' [21]. The questionnaire was answered twice every trial, once before playing to assess the expected level difficulty and once after playing to assess the perceived level difficulty.

Additionally, the changing of affect after every level was measured with the 'Self-Assessment Manikin' [1] containing two visual scales for pleasure and arousal. Finally, the acceptance of technical products was assessed with the German translation and enhancement [8] of the 'Technology-Acceptance-Model' [4] to control that the general acceptance of technical products will not influence the evaluation of the game.

4.2 Procedure

In the beginning of the experiment the participants were welcomed, filled in the confirmation form and were informed about the anonymity of their data, the procedure of the experiment and the possibility to quit at any time. Afterwards, they familiarized with the game and the balance board through a training level (level 7) to avoid influences of familiarity with the board. Then seven trials with the evaluation levels followed in randomized order. In each trial, the participants got a printout of the upcoming level and were asked to answer the pre-questionnaire assessing the difficulty of the level and the first part of the FAM-motivation-questionnaire. Furthermore, if the participants were part of the experimental group, they were asked to set themselves a personal goal. Afterwards they actually played the level until the ball had reached the goal and were informed about the exact time required for achieving this goal. After every level, the participants answered post-questionnaires assessing the difficulty of the level, the second part of the FAM-motivation-questionnaire, the Self-Assessment-Manikin and an evaluation if they reached their goal (if they were in the experimental group). At the end of the session, the participants answered a demographical questionnaire and one on the acceptance of technical products (TAM). Finally, a short interview was done to further evaluate the game. TAM and FAM are well-known in psychological contexts and their reliability has been assessed. Detailed information on that can be found in the corresponding literature.

4.3 Experimental Setup

While playing the different levels the participants stood on the balance board (a product by Japanese manufacturer Nintendo), and able to steer the ball by shifting their weight. The device was positioned in a distance of 0.5 m to a table on which a screen (a Dell Ultrasharp 2009Wt, 20") was set up in 0.9 m height to present the levels. The game BalanceFit was executed on a laptop-computer (a Lenovo Thinkpad T430). The questionnaires were answered while sitting.

All in all, 44 students of psychology and psychology in computer science participated in the study, of which 16 were male and 28 female. Their age varied between 18 and 42 years ($mean$=24.13, SD=5.48). The study was conducted on university students because elderly people have very heterogenous physical abilities which could influence perceived difficulty, motivation and performance, and the aim of this study was to give a more general impression. Surely, another study has to test if the achieved results are also valid for senior citizens.

5. RESULTS AND DISCUSSION

The results were calculated using GNU R [14] Version 3.1.0. The significance level was set to be 5%. Because all tests were calculated within Anovas (Type I), which adjusted the significance levels of the subtests, no adjustment due to multiple comparisons were needed.

Hypothesis 1: Level features influence motivation.

One-way analyses of variance were calculated on the impact of the different level features on motivation. The impact of *level features on pre-game motivation* turned out to be signficant for open curbs, $F(2{,}86)=5.42$, $p=.006^{**}$, $\eta^2=0.11$ and orientation changes, $F(1{,}43)=14.197$, $p<.001^{***}$, $\eta^2=0.25$, but not for bottlenecks, $F(1{,}43)=2.875$, $p=.1$, $\eta=0.063$, or the shortest path, $F(1{,}43)=0.0841$, $p=.77$, $\eta^2=0.002$. There also turned out to be an interaction between the orientation changes and the shortest path, $F(1{,}43)=8.209$, $p=.006^{**}$, $\eta^2=0.16$.

The impact of *level features on post-game motivation* was only signficant for the orientation changes, $F(1{,}43)=9.7646$, $p=.003^{**}$, $\eta^2=0.18$, and again the interaction of orientation changes and the shortest path was signficant too, $F(1{,}43)=10.331$, $p=.003^{**}$, $\eta^2=0.19$. But there was no significance impact of open curbs $F(2{,}86)=1.99$, $p=.14$, $\eta^2=0.044$, bottlenecks, $F(1{,}43)=1.2748$, $p=.27$, $\eta^2=0.029$ or the shortest path $F(1{,}43)=1.031$, $p=.32$, $\eta^2=0.023$.

The length of the open curbs, the orientation changes and the interaction between orientation changes and the shortest path influence the pre-game motivation significantly whereas the number of bottlenecks and the shortest path do not. On post-game motivation only the orientation changes and their interaction with the shortest path seem to have an influence. That would mean, when designing the levels to be highly motivational one should have a close look on orientation changes and their combination with the shortest path.

Hypothesis 2: Level features influence the perceived difficulty.

One-way analyses of variance were calculated on the impact of the different level features on perceived difficulty.

Table 1: Influence of the different level features on the subjectively perceived difficulty before and after playing. The largest possible value was 220, the least was 0.

	Pre-Game		Post-Game	
	mean	SD	mean	SD
Shortest Path				
short (L4,L6)	96.18	51.50	56.55	43.75
medium (L2, L3)	55.56	51.14	29.52	32.12
long (L1,L5,L7)	64.96	44.79	47.86	39.09
Orientation Changes needed				
1 (L4)	94.77	48.94	40.98	29.77
2 (L3)	31.07	37.83	19.36	32.34
3 (L2, L6)	88.82	53.35	55.90	43.67
5 (L7)	52.07	41.36	42.91	39.52
8 (L1)	59.82	40.49	46.48	40.71
9 (L5)	83.00	47.23	54.18	36.99
Bottlenecks				
none (L1,L2,L3,L5,L6,L7)	67.27	50.52	45.79	41.43
At least one (L4)	94.77	48.94	40.98	29.77
Length of open curbs				
none (L2,L3,L5,L7)	61.55	49.25	39.03	36.61
short (L1)	59.82	40.49	46.48	40.71
medium (L4,L6)	96.18	51.50	56.55	43.75

The tests on influence of the *level features on pre-game perceived difficulty* were significant for open curbs, $F(2,86)=$ 25.866, $p<001$***, $\eta^2=.38$, orientation changes, $F(1,43)=$ 38.185, $p<.001$***, $\eta^2=.47$, and the shortest path, $F(1,43)=$ 14.276, $p<.001$***, $\eta^2=.25$, but not for bottlenecks, $F(1,43)=$ 1.150, $p=.29$, $\eta^2=.026$, and again there was a significant result for the interaction of orientation changes and shortest path $F(1,43)=24.468$, $p<.001$***, $\eta^2=.36$.

For the impact of *level features on post-game perceived difficulty* there were significant results for open curbs, $F(2,86)=$ 7.424, $p=.001$**, $\eta^2=.15$, orientation changes ,$F(1,43)=9.7646$, $p<.001$***, $\eta^2=.42$, bottlenecks, $F(1,43)=10.141$, $p<.001$***, $\eta^2=.19$, and again, for the interaction of orientation changes and the shortest path, $F(1,43)=9.5505$, $p=.004$**, $\eta^2=.18$. The results were insignificant the shortest path, $F(1,43)=$ 0.0443, $p=.83$, $\eta^2<.01$.

The perceived difficulty is influenced significantly (directly or indirectly) by all the evaluated features but the bottlenecks, which would mean, that all these features offer the possibility to personalize the level difficulty. An overview over the values of the perceived difficulty of is given in table 1. It is obvious to see that normally, the perceived difficulty is higher before the game than after, but nevertheless, there is a high standard deviation meaning that the difficulty of the levels was perceived very differently by the participants.

Hypothesis 3: Level features influence performance. One-way analyses of variance were calculated on the impact of the different level features on performance.

The impact of *level features on performance* turned out to be significant for open curbs, $F(2,86)=$ 2.323, $p=.01$*, $\eta^2=.05$, orientation changes, $F(1,43)=85.104$, $p<.001$***, $\eta^2=.68$, bottlenecks, $F(1,43)=$ 5.829, $p=.02$*, $\eta^2=.13$ and

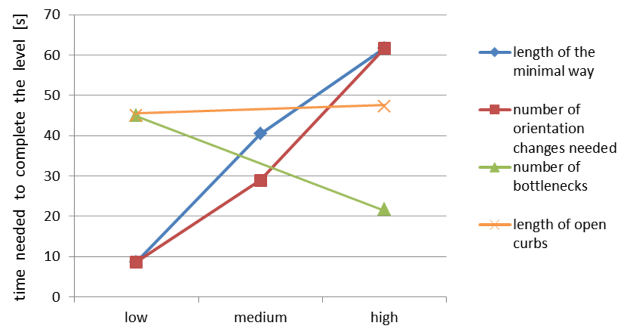

Figure 3: Influence of the level features on the performance.

the shortest path, $F(1,43)=$ 9.098, $p=.004$**, $\eta^2=.18$, and again, there was an interaction effect of orientation changes and the shortest path, $F(1,43)=72.457$, $p<.001$***, $\eta^2=.64$.

That means that all of the chosen level features influence performance significantly, but unexpectedly, there was negative impact of the presence of bottlenecks: If there were bottlenecks, the participants needed less time to complete the level. Maybe the participants felt more insecure and tried to linger less, or maybe there were not enough levels with bottlenecks to draw conclusions.

Hypothesis 4: Goal setting influences motivation, perceived difficulty and performance. One-way analyses of variance were calculated on the impact of goal setting on motivation, perceived difficulty and performance. Unexpectedly, there were no effects of goal setting on neither pre-game motivation, $F(1,42)=0.92$, $p=.34$, $\eta^2=.021$ nor post-game motivation, $F(1,42)=2.54$, $p=.12$, $\eta^2=.057$, pre-game perceived difficulty, $F(1,42)=1.49$, $p=.23$, $\eta^2=.034$, post-game perceived difficulty, $F(1,42)=0.26$, $p=.61$, $\eta^2=.006$, or performance, $F(1,42)=0.6873$, $p=.41$, $\eta^2=.02$.

For the influence of goal setting, none of our results turned out to be significant. That may be due to the fact that there is no effect of goal setting, but it is more likely that one of the reasons Locke and Latham [9] listed may have interfered: In our case, it is possible that the variation of the level features was not consistent enough or we did not include a sufficient range of difficulty levels. Alternatively, we might not have gotten goal commitment.

6. SUMMARY AND OUTLOOK

In this paper the influence of individual game features on the perceived difficulty posed on the player have been investigated. The results show a significant influence of single features on psychological (motivation, difficulty) and physiological constructs (performance, needed time).

The configuration of the level features can influence the amount of time needed to succeed and the motivation to play. Therefore, it may be reasonable to present levels with shorter shortest paths when the participant begins the training and increase the length successively. The number of orientation changes should also increase with the participants ability to play the game, because shortest path and number of orientation changes both seem to influence the motivation positively.

In order to allow more fine-grained studies, a system for the creation of levels considering the evaluated features is developed. This system allows to build levels manually and automatically and takes into consideration the findings of this study.

7. ACKNOWLEDGMENTS

This research was partially financed by the European Commission under the FP7-ICT Project ALFRED (grant agreement no. 611218). The authors would like to thank Nina Kolb, Verena Schochlow, Janine Schwan, and Anna-Livia Meimbresse, and Lyenne Spieske who supported the the data acquisition for the study presented in this paper.

8. REFERENCES

[1] M. M. Bradley and P. J. Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994.

[2] S. Consolvo, P. Klasnja, D. W. McDonald, and J. A. Landay. Goal-setting considerations for persuasive technologies that encourage physical activity. In *Proceedings of the 4th international Conference on Persuasive Technology*, page 8. ACM, 2009.

[3] M. Csikszentmihalyi. *Das flow-Erlebnis: jenseits von Angst und Langeweile: im Tun aufgehen*. Klett-Cotta, 1985.

[4] F. D. Davis Jr. *A technology acceptance model for empirically testing new end-user information systems: Theory and results*. PhD thesis, Massachusetts Institute of Technology, 1986.

[5] K. Eilers, F. Nachreiner, and K. Hänecke. Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. *Zeitschrift für Arbeitswissenschaft*, (4):214–224, 1986.

[6] S. Hardy, S. Göbel, M. Gutjahr, J. Wiemeyer, and R. Steinmetz. Adaptation model for indoor exergames. In *International Journal of Computer Science in Sport Vol. 10, Special Edition: Serious Games - Theory, Technology and Practice*, 2011.

[7] S. Hardy, S. Göbel, and R. Steinmetz. Adaptable and personalized game-based training system for fall prevention. In *ACM Multimedia*, pages 431–432, 2013.

[8] M. Kauer. Identification as determining factor of technology acceptance for hedonic and dual use products. 2012.

[9] E. A. Locke and G. P. Latham. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, 57(9):705, 2002.

[10] J. Maitland and K. A. Siek. Technological approaches to promoting physical activity. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, pages 277–280. ACM, 2009.

[11] F. Mehm, S. Hardy, S. Göbel, and R. Steinmetz. Collaborative authoring of serious games for health. In *Proceedings of the 19th ACM International Conference on Multimedia*, MM '11, pages 807–808, New York, NY, USA, 2011. ACM.

[12] W. Press, editor. *The World Health Report 2013*. World Health Organization, 2013.

[13] A. K. Przybylski, C. S. Rigby, and R. M. Ryan. A motivational model of video game engagement. *Review of General Psychology*, 14(2):154, 2010.

[14] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014.

[15] F. Rheinberg, R. Vollmeyer, and B. D. Burns. FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern-und Leistungssituationen. *Diagnostica*, 47(2):57–66, 2001.

[16] R. M. Ryan and E. L. Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1):68, 2000.

[17] J. Sinclair, P. Hingston, and M. Masek. Considerations for the design of exergames. In *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia*, GRAPHITE '07, pages 289–295, New York, NY, USA, 2007. ACM.

[18] P. Sweetser and P. Wyeth. Gameflow: A model for evaluating player enjoyment in games. *Comput. Entertain.*, 3(3):3–3, July 2005.

[19] M. E. Tinetti. Performance-oriented assessment of mobility problems in elderly patients. *Journal of the American Geriatrics Society*, 1986.

[20] R. J. Vallerand and G. F. Losier. An integrative analysis of intrinsic and extrinsic motivation in sport. *Journal of applied sport psychology*, 11(1):142–169, 1999.

[21] F. Zijlstra and L. v. Doorn. The construction of a scale to measure subjective effort. *Delft, Netherlands*, 1985.