How to Evaluate Serious Games Concepts: A Systematic Prototyping and Testing Approach

Cornelia Schade and Antonia Stagge
Technische Universität Dresden, CODIP, Germany
cornelia.schade@tu-dresden.de
antonia.stagge@tu-dresden.de

Abstract: The challenge in developing a serious game is to find the perfect balance between learning and playing. The development process should include an appropriate involvement of the target group and enable a systematic evaluation of this balance through prototyping and testing. The goal is to create an entertaining and purposeful learning experience and thus enable knowledge growth. This paper presents the evaluation results of the serious game E.F.A. with the target group – managers in the social service sector. The first prototype was tested in an early phase as a paper prototype by experts in media didactics and subject experts. Early stage testing is a decisive factor for the development of serious games. However, the accessibility of the target group is not always given for fast testing and iterative improvement. After collecting expert feedback and incorporating it into the game, the high-fidelity prototype was created and tested by the target group. Those test runs were followed by group interviews. Their results are the focus of this paper which aims at answering the following research questions: How did the target group experience the serious game and their increase in knowledge? To what extent can the evaluation results with the target group be linked to the early tests with the paper prototype? How does the feedback vary and what conclusions can be drawn from this? The results of the paper show that the serious game was rated very differently among the target group. Some generally praised the playful approach. Others criticized the game as childish and unsuitable for the target group. The feedback obtained from different user groups with the help of different prototypes varied for a set of evaluation criteria such as playing time, remembered knowledge and dialogs. For each evaluation criteria recommendations are given regarding the test group and type of prototype.

Keywords: serious games, evaluation, user-oriented, learning experience design, prototyping

1. Prototyping and testing during the development of the serious game E.F.A

When developing serious games, a balance between learning and entertainment must be established. This means achieving the intended learning objectives while creating enjoyable experiences (Marr 2010, Lampert et al. 2009). This can increase the motivation of the learners (Egenfeldt-Nielsen 2005). Within ideation, prototyping and testing phases of serious games different professional perspectives need to be combined. The concept of learning experience design (LXD) takes up this interdisciplinary approach. The main focus of LXD is to create “more engaging, innovative and effective experiences for learners in different environments” (Ahn 2018, p. 1). This young design discipline combines specific design knowledge and methods (from interaction design, user experience design, human-centered design, etc.) with experience from the field of methodological-didactic design of learning experiences and from IT/software development, among others (Kircher et al. 2021). The increasing collaboration between different disciplines as well as the rising costs for creating digital learning materials like serious games, require a refinement and optimization of development processes. It is becoming more important that solutions are developed and tested as early and effectively as possible so a large part of the resources can go into creating the final design (Cennamo, Kalk 2019). For that, the target group should be involved as early as possible to ensure that the game meets their needs. This can be achieved by testing the game concept using prototypes during the development process to obtain feedback from the user group (Ampatzidou and Gugerell 2019, Sailer 2016).

In order to examine the development process of serious games this paper focuses on the project E.F.A. which aims at developing a serious game for the purpose of improving occupational safety and health (OSH) and its implementation, especially in small enterprises. The player takes on the role of an expert in OSH and runs through four temples in the "obligation jungle". The most important aspects of OSH are imparted through various learning tasks (mini-games) and a knowledge repository. Inserted dialogues convey the game story.

During the development of the serious game E.F.A. a systematic approach was applied in which phases of prototyping and concept development alternated with test phases (see figure 1). As soon as the first didactic concept was available a prototype was developed that enabled the testing of the game.

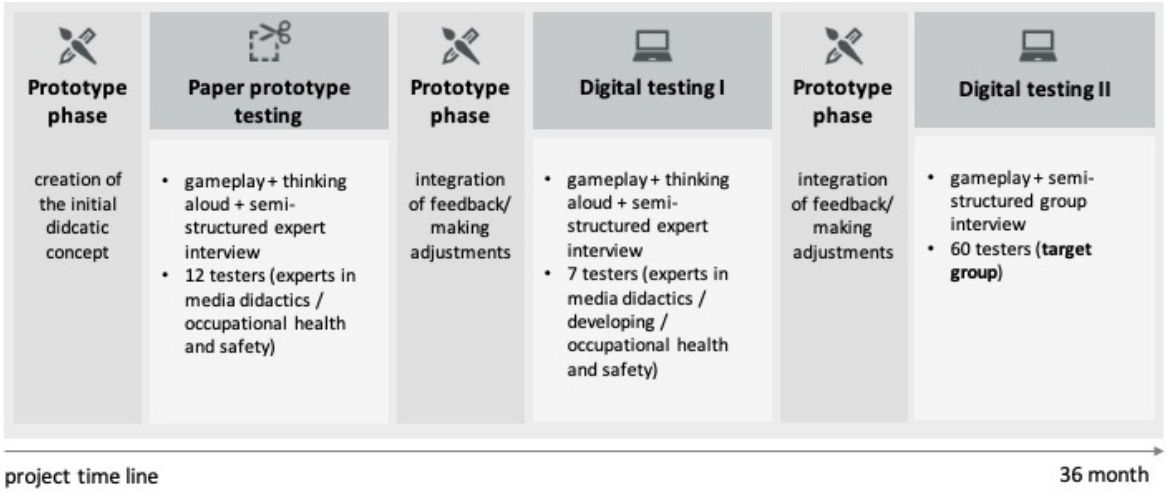

Figure 1: Systematic prototyping and testing scheme for the serious game E.F.A

A distinction can be made between low and high fidelity prototypes. A detailed overview of the classification of fidelity dimensions can be found in Zdralek (2000). Low-fidelity prototypes are used to quickly and without great cost generate feedback for an idea in early development phases (Abidin et al. 2019, Giacomelli et al. 2018). Paper prototypes are a specific form of low-fidelity prototypes. They can be described as "a variation of usability testing where representative users perform realistic tasks by interacting with a paper version of the interface that is manipulated by a person “playing computer” [...]" (Snyder 2003, p. 4). They are also used in game development where they function as a kind of board game. The logic and game mechanics are often similar to those of the digital game that is going to be developed (Schell 2016). High-fidelity prototypes are already very close to the final product and more technically sophisticated. However, they require much more time and resources (Zhang et al. 2012). There is still disagreement in the scientific discourse about the effective use, limitations and validity of various forms of prototypes during different stages of development (e.g. Wiklund et al. 1992, Zhang et al. 2012, Sauer and Sonderegger 2009, Sefelin et al. 2003, Walker et al. 2002, Virzi et al. 1996). But it seems clear that "effective and efficient evaluation methods and measurements are essential to the success of the iterative design process" (Zhang et al. 2012, p. 2).

By implementing a test phase in the E.F.A. project at an early stage, the game concept could be comprehensively evaluated with the simplest possible means before it was transferred to a digital version. Therefore, the overall prototyping and testing scheme used different forms of prototypes with different levels of fidelity. In a first phase the paper prototype was extensively tested with experts from various disciplines. The expert feedback was incorporated into the didactic concept and the first digitally playable prototype was created. This was again tested with different expert groups. A comparison of these two test phases already showed that the paper prototype was a useful instrument for testing basic elements of a serious game, such as the complexity and suitability of the learning tasks. In order to test more specific elements and the usability the digitally implemented version was more suitable (Stagge and Schade 2021). After the first two test phases were completed, the digital prototype was finalized and tested with the actual target group.

An efficient and effective development process involves not only choosing the right prototype at the right time, but also paying attention to the actual test participants and the type of feedback they can generate. For early tests with rather immature prototypes, it is advantageous to choose a test group with good access from which benevolent feedback can be expected. Once the prototype has been further developed, test persons who are more distant from the development team and who are not familiar with the prototype should be involved (Kapp et al. 2013). E.F.A. was designed for the target group of the social services sector. This includes hard-to-reach groups, such as nursing. Therefore, the full implementation of a user-centered design process was a great challenge. At the beginning of the project the target group and the context of use were analyzed comprehensively. For the early testing, the team mainly relied on experts from different disciplines, as they were easily accessible and thus a fast cycle of testing and adapting could be established. This approach raised the question of the extent to which early prototypes for serious games can also be tested by a wider circle of experts and still result in a product that is considered useful by the target group.

2. Research Questions

This paper addresses the following research questions:

- How did the target group experience the serious game and their increase in knowledge?
- To what extent can the evaluation results with the target group be linked to the early tests with the paper prototype? How does the feedback vary and what conclusions can be drawn from this?

By answering these two research questions, the paper reveals how E.F.A. was perceived and evaluated by the target group regarding a pre-defined set of evaluation criteria (section 3). These results paired with the already described systematic testing scheme (section 1) provide practical insights that can be used by future designers of serious games.

3. Methodology

An iterative testing approach was used in the evaluation of E.F.A.. The first draft of a didactic concept resulted in a paper prototype, which was modified after each test run according to the given feedback. The digital prototype was also first tested with experts (see comparison of both test runs in Stagge and Schade 2021) before being evaluated with the target group. This paper presents the results of the tests with the target group. It also compares the results of the paper prototype tests by the expert groups with the results of the digital prototype tests by the target group (see figure 2).

Participants. The digital prototype was tested by 60 persons of the target group. The paper prototype was tested by eleven experts from the field of media didactics and by one expert from the field of OSH. Their expertise enables them to evaluate the comprehensibility of the didactic approach and the game content regarding the relevant subject knowledge.

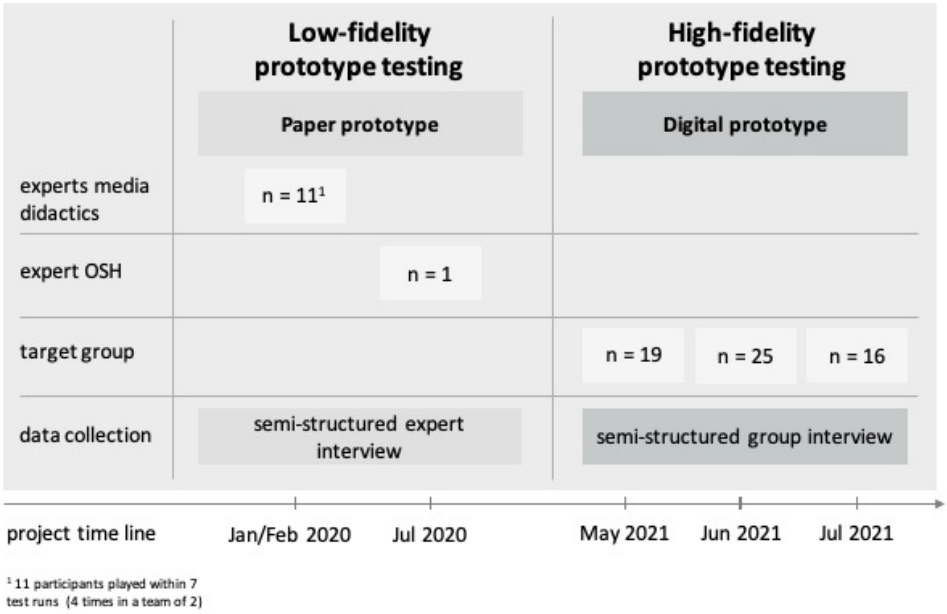

Figure 2: Overview of the test runs

Test procedure. The test runs with the digital prototype were integrated into an online test event in which participants had 45 minutes to individually test the prototype. Afterwards participants were divided into groups and interviewed about their game impressions and perceived knowledge gain using three questions in a semi-structured interview.

The testers of the paper prototype were asked to think aloud (Abidin et al. 2019) throughout playing the game. In the subsequent semi-structured interview feedback on specific evaluation criteria was obtained (Schade and Stagge 2020).

Data collection and analysis. Data was collected through the semi-structured interviews. The testers' answers were recorded and the data was analyzed for content with deductive category formation. The following 15 categories were defined beforehand:

1 First Impression	6 Perceived Increase in Knowledge	11 Fit for Target Group
2 Playing Time	7 Remembered Knowledge Content	12 Negative Aspects
3 Design/Visuals	8 Game Control	13 Positive Aspects
4 Mini-Games	9 Usability/User Guidance	14 Proposals for Further Development
5 Task Complexity	10 Dialogs	15 Transfer of Learning into the Company

The answers to the questions were collected and assigned to the categories. The results of the test runs allow conclusions about how the target group perceived the serious game and their increase in knowledge.

Then these results were compared with those from the experts' test runs with the paper prototype (for the results of the paper prototype tests see Schade and Stagge 2020). Here only nine of the 15 evaluation criteria (1, 2, 4, 5, 7, 10, 12, 13, 15) of the digital prototype could be used for comparison purposes. The reason lies in the differently designed interview questions. In the paper prototype test, a total of ten questions were asked that targeted specific evaluation criteria. In the test with the target group, three more open-ended questions were used due to the group setting. By comparing the data, differences and similarities in the evaluation of the serious game could be identified.

4. Results

The following section presents the results of the test runs with the target group (section 4.1). Subsequently, these results are related to the evaluation of the paper prototype (section 4.2).

4.1 Evaluation of the digital prototype by the target group

1 First impression. It was very mixed: from boring, exhausting, quite okay to entertaining and easy to understand. The jungle setting, eye-catching hazards and puzzles were praised, as was the game's approach in general – even if it caused frustration for some people.

2 Playing time. It was evaluated very differently. Some felt it was brisk. Others felt it was quite long, especially due to the long walking distances. Others – especially older or unexperienced players – felt under time pressure and would have needed more time to fully explore the game probably due to the test setting.

3 Design/Visuals. Some found it very childish, others found it well designed. Especially the color scheme was praised. The visual treatment was very memorable: the very figurative language contributed to an easy understanding.

4 Mini-Games. There was little feedback on mini-games. However, some praised them as a good implementation for entering/exiting the temple. The short knowledge repository displayed at the end of each mini-game, as well as the link to the more detailed knowledge repository, also received positive reviews.

5 Task complexity. Some had problems with the tasks inside the temple (pitfall, extinguishing the fire) and would have liked a short tutorial. A forgotten photo documentation (one of the mission goals) prevented some from easily traversing the temple. Otherwise, most testers found the tasks quite simple and would have liked even greater challenges.

6 Perceived increase in knowledge. Most testers were able to take away some knowledge, even if it was mainly a repetition of pre-existing knowledge. The game acted as an impulse to take a closer look at the topic. Others estimated that the increase in knowledge was not sustainable, especially if the topic of OSH is not a priority in their daily work life. Some users focused more on playing and less on content.

7 Remembered knowledge content. Especially those contents could be remembered where the players had to become active themselves or which were implemented very pictorially, e.g. fire protection, first aiders, or documentation of accidents. Very specific content could also be recalled, for example facts about insurance coverage on the commute to work.

8 Game control. The controls of the game character were predominantly criticized and took time to get used to, which made solving the tasks in the temple more difficult. This was especially hard when playing on a mobile device. Switching between keyboard and mouse was also felt to be annoying.

9 Usability. The main criticism was the lack of orientation within the game environment. The execution of actions in the temple was not intuitively designed. Interaction areas were sometimes not recognizable. On mobile devices, the display of some game content was too small for some participants.

10 Dialogs. The dialogs were often criticized for their large amount of text. They were perceived as funny but also as too childish. Especially the character of the parrot seemed too dominant.

11 Fit for the target group. Some players didn't find the game appropriate, but very childish and not challenging enough. They felt it was more suitable for children/young people and possibly too complex for older people. Others found it very suitable for their own daily work.

12 Negative aspects. The many interruptions and thus the many necessary clicks (e.g. through dialogs) stood out as well as inappropriate sound effects (e.g. exuberant applause). The chosen metaphor on the subject of fatal accidents was felt to be too inappropriate for such a serious topic.

13 Positive aspects. Many liked the idea of a playful implementation to lighten up the theoretical content. It was also praised that the information was conveyed in an understandable and reduced manner. Some sound effects and the possibility to control the game character had a motivating effect. The character of the parrot as a companion and the knowledge repository were also well received.

14 Proposals for further development. Most recommendations for further development related to game play actions and user-friendliness: Testers wanted more action, animation, sounds, and ways to interact with the character (e.g. jumping), as well as more tutorial guidance on completing actions. Additionally, the testers would like to see more references to their everyday working life and more detailed information.

15 Transfer of learning into the company. The knowledge repository was considered very relevant – even more relevant than the game itself. Also the documentary function supported the knowledge transfer. The "very striking" hazards in the game were felt to be easily transferable to the testers' everyday working lives. Some testers felt that the added value was too low and that the game is too unattractive to use it in practice.

4.2 Evaluation of the paper prototype with experts vs. digital prototype with the target group

1 First impression. The game play was perceived as entertaining and fun by most experts. The target group gave mixed feedback – from boring and exhausting to interesting and easy to understand. Both test phases showed that the players needed some time to become familiar with the game.

2 Playing time. During paper prototype testing the playing time was generally perceived as entertaining and optimal. The digital prototype test revealed a very mixed rating: from long to brisk.

4 Mini-games. The mini-games were perceived positively by the experts but the implementation was often criticized, especially their ability to convey knowledge. This feedback prompted a complete revision of the didactic concept of the learning tasks. The target group praised the mini-games as a good implementation.

5 Task complexity. For some testers of the paper prototype, the tasks were easy, others had problems with orientation and the high amount of information. In digital testing it was two-folded: the complexity of learning tasks was perceived as optimal to too easy and game play tasks (e.g. crossing the pitfall) were perceived as difficult to perform.

7 Remembered knowledge content. In the paper-based test little content stuck with most testers – probably due to the initial insufficient effectiveness of the mini-games. The digital tests showed that the players could remember many topics, technical terms and concrete facts.

10 Dialogs. The dialogs were largely perceived positively by the experts, with praise for their looseness and wit. They were also rated as helpful by the OSH expert. Even though the target group perceived the dialogs as funny, they were mainly criticized for their large amount of text and childish character.

12 Negative aspects. Long dialogs and thus the many interruptions were cited as the greatest point of criticism by media didactic experts and the target group. The media didactic experts also rated running back when the documentation was forgotten as demotivating. Some had difficulties with the mini-games as well as with the transfer from game to reality. The OSH expert mainly had content-related criticism (e.g. missing content, more hazard examples). The target group was mainly irritated by some sound effects or metaphors chosen to depict different topics.

13 Positive aspects. The OSH expert and the target group alike especially praised the short knowledge repository as a useful addition. The media didactic experts rated the dialogs, characters and the ambience of the game as positive. The target group praised the animal companion, as well as the idea of the playful implementation of a theoretical topic. They also positively evaluated the possibility to actively move through the game environment.

15 Transfer of learning into the company. All test groups alike rated the documentation as supportive for the transfer into reality. However, the OSH expert criticized that the hazards should have a higher relevance for the target group and a stronger relation to their work routine. A large part of the target group, on the other hand, felt that the "very striking" hazards were easily transferable. Some also praised the knowledge repository. Others felt that the added value of the game was too low for use in practice.

5. Conclusion

The serious game was rated very differently among the target group. While some emphasized the playful approach positively, could handle the mini-games, found the task difficulty easy or even too easy and were able to retrieve knowledge, others were less enthusiastic. They found the game childish and not suitable for the target group, had problems especially with gameplay actions and were not really able to take away knowledge.

In comparison with the paper prototype tests many differences but also some similarities were found with regard to the evaluation criteria. Variations in feedback existed primarily for the evaluation criteria first impression, playing time, remembered knowledge content, dialogs and transfer of learning into the company (criterion 1, 2, 7, 10 and 15).

It is to be examined in more detail below which evaluation criterion can be tested best with which test group. In addition, conclusions can be drawn from the previous paper (Stagge and Schade 2021) about which type of prototype is best suited for this purpose. Table 1 summarizes these conclusions.

Table 1: Conclusions and recommendations for different evaluation criteria based on the serious game E.F.A.

Criterion	Conclusions and recommendations
First Impression	• Feedback from the target group varied from that of the experts • Expert groups can give a first indication of the first impression of a game *Test with the target group at an early stage*
Playing time	• Feedback from the target group varied from that of the experts • Paper prototype only gave an approximate impression of how the playtime was perceived *Test with a high-fidelity prototype together with the target group*
Mini-games	• Experts can sufficiently test the didactical concept and implementation of the relevant content • Valuable feedback for further conception was obtained here – already with a paper prototype and without the target group *Test with experts at an early stage, iteratively improve concept and check with the target group*
Task Complexity	• Experts can already evaluate the general task complexity • For gameplay actions within the game environment, the paper prototype was not suitable *Test with experts at an early stage (for gameplay actions with a higher fidelity prototype), iteratively improve concept and check with the target group*

Remembered Knowledge Content	• Experts have already identified major weaknesses in the transfer of knowledge with the help of the paper prototype *Test with experts at an early stage, iteratively improve concept and check with the target group*
Dialogs	• Feedback from the target group varied from that of the experts *Test with the target group at an early stage*
Negative / Positive Aspects	• Experts already gathered many positive and negative aspects by means of the paper prototype • Feedback from the target group varied in some aspects *Test with the target group at an early stage*
Transfer of Learning into the company	• Experts were able to make an initial assessment, especially the OSH expert • Target group provided a more comprehensive assessment *Test with the target group at an early stage*

It can be concluded that for evaluation criteria 1, 2, 10, 12, 13 and 15, it is recommended to test directly with the target group. In principle, both prototypes are suitable for testing these criteria, except for criteria 2 and 5. Here a high-fidelity prototype is recommended for evaluation. Since simple changes and improvements can only be made to low-fidelity prototypes without great effort, testing with this prototype at an early stage of development should be considered.

6. Discussion and future research

This paper shows that feedback of different user groups of the serious game E.F.A. using different prototypes varies for certain evaluation criteria. Previous research already gave similar indications. For example, Sellen et al. (2009) claimed that prototype formats and tests groups have an effect on the generated feedback. In the literature on human-centered design, a distinction is made between different groups that test as part of the evaluation — for example, experts in a particular field or the target group (Nelles et al. 2018). There, however, the focus is usually exclusively on usability, which was only one evaluation criterion within this study. This work relates to the early testing of serious game concepts and includes a broader range of evaluation criteria. This is in line with the LXD approach which focuses equally on aspects such as usability and interface design as well as on learning experience. In addition, the paper extends the existing research by the aspect of distinguishing between different test groups of a serious game. In Schade and Stagge 2020 and Stagge and Schade 2021 it was already evaluated to which extent different types of prototypes are sufficient for testing serious game concepts. This study now expands previous research by taking a closer look at the test groups while including previous findings on the type of prototypes.

However, the results of the paper stem from a single case study only. Besides, the sample size in the presented assessments was quite small. However, to achieve a comprehensive understanding of the end-users Maheu-Cadotte et al. recommended to test with a small number of the target group, going more in-depth than broad by using qualitative methods such as interviews and focus group discussions. This is suggested already during the early development phase (Maheu-Cadotte et al., 2021). Moreover, the group interview setting used in this study must be viewed critically. Given answers were often adapted to those of previous speakers and possibly answered according to social desirability. Thus, the answers may not be as differentiated as they would have been in individual interviews. Also, during content analysis, when summarizing statements, there is a risk of biased evaluation or over-interpretation of results.

References

Abidin, S. R. Z., Noor, S. F. M. and Ashaari, N. S. (2019) “Low-fidelity Prototype Design for Serious Game for Slow-reading Students”, International Journal of Advanced Computer Science and Applications (IJACSA), Vol 10, No. 3, pp 270–276.

Ahn, J. (2018) “Drawing Inspiration for Learning Experience Design (LX) from Diverse Perspectives”, The Emerging Learning Design Journal, Vol 6, No. 1, pp 1–6.

Ampatzidou, C. and Gugerell, K. (2019) “Participatory game prototyping – balancing domain content and playability in a serious game design for the energy transition”, CoDesign, Vol 15, No. 4, pp 345–360.

Cennamo, K. and Kalk, D. (2018) Real World Instructional Design. An Iterative Approach to Designing Learning Experiences, Routledge, New York.

Egenfeldt-Nielsen, S. (2005) Beyond Edutainment: Exploring the Educational Potential of Computer Games. IT-University Copenhagen.

Giacomelli, J. M., Fernandes, K. C., Lopes, G. M., de Souza Moura, J., Patikowski Cheiran, J. F. and de Mello, A. V. (2018) “Game Design for Training of Elderly People in Mouse and Keyboard Use”, SBC - Proceedings of SBGames, pp 1345–1348.

Kapp, K. M., Blair, L. and Mesch, R. (2013) The gamification of learning and instruction fieldbook: Ideas into practice, John Wiley and Sons, San Francisco, CA, USA.

Kircher, J., Burger, E.-M., Ebner, M. and Schön, S. (2021) "Learning Experience Design - zur Gestaltung von technologiegestützten Lernerfahrungen mit Methoden der Design-Entwicklung", K. Wilbers and A. Hohenstein (Eds.), Handbuch ELearning. Expertenwissen aus Wissenschaft und Praxis – Strategien, Instrumente, Fallstudien, Bd. Erg.-Lfg. 93, 4.78.

Lampert, C., Schwinge, C. and Tolks, D. (2009) "Der gespielte Ernst des Lebens: Bestandsaufnahme und Potenziale von Serious Games (for Health).", [online], Medienpädagogik: Zeitschrift für Theorie und Praxis der Medienbildung. Themenheft, No. 15/16, https://www.medienpaed.com/article/view/104.

Maheu-Cadotte M., Dubé V., Cossette S., Lapierre A., Fontaine G., Deschênes M. and Lavoie P. (2021) "Involvement of End Users in the Development of Serious Games for Health Care Professions Education: Systematic Descriptive Review", JMIR Serious Games, Vol 9, No. 3.

Marr, A. C. (2010) "Serious Games für die Informations- und Wissensvermittlung. Bibliotheken auf neuen Wegen", B.I.T.online – Innovativ, Vol 28, Dinges & Frick GmbH, Wiesbaden.

Nelles, J., Brandl, C., Mertens, A. (2018) "Regelkreismodell für die menschzentrierte Gestaltung und Evaluierung einer Mensch-Roboter-Interaktion am Beispiel eines Mensch-Roboter-Arbeitsplatzes", Arbeit(s)Wissenschaf(f)t – Grundlage für Management und Kompetenzentwicklung. 64. Kongress der Gesellschaft für Arbeitswissenschaft, pp 1–6.

Sailer, M. (2016) Die Wirkung von Gamification auf Motivation und Leistung: Empirische Studien im Kontext manueller Arbeitsprozesse, Springer, Wiesbaden, Germany.

Sauer, J. and Sonderegger, A. (2009) "The influence of prototype fidelity and aesthetics of design in usability tests: Effects on user behaviour, subjective evaluation and emotion", Applied Ergonomics, Vol 40, No. 4, pp 670–677.

Schade, C. and Stagge, A. (2020) "Paper Prototyping as a Method for the Evaluation of Serious game Concepts", European Conference on Games Based Learning, Academic Conferences International Limited, pp 515–521.

Schell, J. (2016) Die Kunst des Game Designs: Bessere Games konzipieren und entwickeln, mitp, Frechen, Germany.

Sefelin, R., Tscheligi, M. and Giller, V. (2003) "Paper prototyping-what is it good for? A comparison of paper-and computer-based low-fidelity prototyping", CHI'03 extended abstracts on Human factors in computing systems, pp 778–779.

Sellen, K. M., Massimi, M. A., Lottridge, D. M., Truong, K. N. and Bittle, S. A. (2009) "The people-prototype problem: understanding the interaction between prototype format and user group", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp 635–638.

Snyder, C. (2003) Paper prototyping: The fast and easy way to design and refine user interfaces, Morgan Kaufmann, Amsterdam, Netherlands.

Stagge, A. and Schade, C. (2021) "Paper-based vs. Digital Prototyping: How to Evaluate Serious Game Concepts at Different Stages of Development", European Conference on Games Based Learning, Academic Conferences International Limited, pp 692–699.

Virzi, R. A., Sokolov, J. L., and Karis, D. (1996) "Usability problem identification using both low- and high-fidelity prototypes", M. J. Tauber, V. Bellotti, R. Jeffries, J. D. Mackinlay and J. Nielsen (Eds.), Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (SIGCHI'96), ACM Press, New York, NY, pp 236–243.

Walker, M., Takayama, L. and Landay, J. A. (2002) "High-Fidelity or Low-Fidelity, Paper or Computer? Choosing Attributes when Testing Web Prototypes", Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol 46, No. 5, pp 661–665.

Wiklund, M. E., Thurrott, C. and Dumas, J. S. (1992) "Does the fidelity of software prototypes affect the perception of usability?", Proceedings of the Human Factors Society Annual Meeting, Vol 36, No. 4, pp 399–403.

Zdralek, J. F. (2000) "Clarifying the Fidelity Dimensions of Prototypes", [online], http://www.deziner.com/Papers_files/Prototyping_Fidelity.pdf.

Zhang, T., Patrick Rau, P. L., Salvendy, G. and Zhou, J. (2012) "Comparing Low and High-Fidelity Prototypes in Mobile Phone Evaluation", International Journal of Technology Diffusion (IJTD), Vol 3, No. 4, pp 1–19.