Designing an Engaging and Informative Application About First Aid: Gamification and Humor as Design Elements in a Serious Game

Nicolai Foldager, Hans Hansen, Mikkel Skovsmose Tewes, and Thomas Bjørner(✉)

Department of Architecture, Design and Media Technology, Aalborg University, A.C. Meyers Vænge 15, 2450 Copenhagen, SV, Denmark
tbj@create.aau.dk

Abstract. This study aimed at developing an engaging and informative application within first aid and CPR for people who are already certified in first aid. The paper outlines discussions within definitions of serious games, humor, gamification and engagement. Further we suggest specific elements for implementation and evaluation of humor and gamified elements. Two prototypes were developed: one with gamification elements and one without. A between-group design was used, in which two different groups tested one prototype each. Data were gathered through data logging, in-depth interviews (with use of a verbal numeric rating scale) and observations of participants' facial expression. The Facial Action Coding System (FACS) was used for analysis. The results showed very little difference between the gamified and non-gamified version. Important elements within gamification are focuses and thoughtfulness within the implementation of challenge, rewards, achievements, feedback and the overall visual theme.

Keywords: Serious game · Gamification · Enjoyment · Humor · Engagement · Qualitative · Observations

1 Introduction

Out-of-hospital cardiac arrest is a major health problem associated with poor outcomes [1, 2]. Every year, 3,500 people experience an out-of-hospital cardiac arrest in Denmark, and only 57.9% receive cardiopulmonary resuscitation (CPR) before an ambulance arrives. The number of CPR recipients has increased over the last 10 years [1], which could be due to a focus on supportive ubiquitous technology solutions, e.g., the large amount of applications as life-saving tools and use of GPS for exact location and nearby automated external defibrillator (AED) heart starters. Training courses with first aid (and CPR) are also mandatory in Danish state schools and are a condition for obtaining a driver license. The number of people receiving CPR can, however, still increase,

O. Gaggi et al. (Eds.): GOODTECHS 2016, LNICST 195, pp. 78–87, 2017.
DOI: 10.1007/978-3-319-61949-1_9

considering that receiving CPR triples the survival rate [1, 2] and improves long-term outcomes [1, 2]. One important factor (and sometimes one that is overlooked) is to focus on keeping the already-trained layman up-to-date with CPR knowledge. This paper presents a study in which a mobile application is developed with a focus on different game design elements. The aim is to motivate people who have already been trained in CPR to refresh their first-aid knowledge using a serious game. The research question is as follows: how can game design elements be used to develop an engaging and informative application about first aid for people who are already trained? Previous research had similar approach [28, 29], however one element in this paper is also to evaluate gamification elements and humor to gain the objectives of an engaging and informative application.

There is no consensus on the definition of serious games, and they are used in divergent ways, focusing on various perspectives within purposes, players'/users' goals, and intended content [7, 23]. Some general requirements embedded in serious games are suggested as follows [23]. 1. The game play is intrinsically motivating. 2. There is immediate feedback in the game environment. 3. The content can have or has learning opportunities. One suggested definition of serious games that might also work as a starting point is "Any form of interactive computer-based game software for one or multiple players to be used on any platform and that has been developed with the intention to be more than entertainment" [23, p. 6]. However, within this definition, there still might be some unsolved categorical problems of what "a game" is and what "entertainment" actually means. The main problem is still that the various definitions of serious games are too generic, and we would rather use the term gamification. Though there are many different perspectives within gamification, there is, however, some agreement that gamification can be seen as the process of game-thinking and game mechanics to engage users to solve problems within a non-game context [3, 4]. The purpose of using gamification elements is improving the user experience, which includes engagement. Engagement is defined as "the desire to continue playing" [5] which can be seen as a prerequisite for the experience of other conceptualizations such as engrossment, flow, fun, enjoyment, immersion, involvement, and incorporation because a player first needs to desire continue playing before these other aspects of the player experience can be experienced [5]. There is significant hype around gamification, and even though there are some good examples, there are also great risks for failure. This can be due to poor design and a lack of knowledge about the target group. Careful thought is also necessary about both the advantages and disadvantages of the design methods.

2 The Design as Engaging and Informative

Scholars have come up with different definitions of engagement and suggestions for increasing players' engagement [5, 8–10]. There seems to be some agreement on the importance of motivation, challenge, and flow. O'Brien and Toms [9] suggest that engagement shares attributes such as intrinsic motivation with flow. Motivation is needed to interest users in the activity and is what makes a user enjoy an activity [13]. Challenge is an important attribute for engagement [9, 11] and is also described as flow [13, 14], in which the player's skills and challenge must be closely matched [14, 15].

Based on the theoretical framework, especially that of O'Brien and Toms [9] and Sweetser and Wyeth [11], the game design was developed with the following criteria:

1. Several scenarios. The application should consist of several different scenarios to maintain users' interest and motivation and to gain specific and different learning goals within CPR. As two-thirds of cardiac arrest cases occur at home, it was a prerequisite that a "home scenario" was implemented in different situations, e.g., in terms of different numbers of bystanders (0–3).
2. Implementation of humor. Because the target group for this application is young individuals (20–28 years of age) different elements involving humor were implemented. In the literature, it is described how humor attracts attention and can be used to make a dull subject more interesting [17–19] and even enhance learning [20]. Because first aid is a subject that can be quite tiresome to read about, we designed some of the answering options in the challenges to have humoristic formulations. This should help users obtain a certain amount of enjoyment, which is a part of the positive effect [9, 11]. All users are not motivated by the same things; some prefer hard fun, and others prefer easy fun [18]. It is therefore crucial to have in-depth qualitative user insight for implementing humor as intended.
3. Implementation of rewards in the application. Rewards can give the user clear visual feedback based on performance, which can motivate the user to try the scenarios more than once [9, 11, 12]. Furthermore, the users will be unable to continue to scenario 2 [see Fig. 1 for example] if they received less than two stars in the first scenario. Stars were chosen to indicate points because they are an already well-known reward symbol (as known from, e.g., Angry Birds).
4. Implementation of achievements. Achievements can be a motivational factor [9] because they provide clear feedback when showing users their progress. They also give users a reason to continue using the application (meaning they are engaged [5]), especially if they want to unlock as many scenarios as possible [see Fig. 1 for example]. The achievements will be shown as highlighted with a bright yellow color to give a clear indication of the user's accomplishments. The user will be shown a popup that enters the screen from the top to give them instant feedback as an important element in game flow [9, 11].
5. Implementation of challenge. It is important that the difficulty level of the content in the application is higher than the user's skill level [9]. The challenge will be implemented by different questions about CPR.
6. Implementation of feedback. A horizontal progress bar was implemented that fills up based on the user's progress (see Fig. 1), similar to how hit point bars and energy bars usually are displayed in games. Furthermore, the color of the progress bar was set to be red because it would match the rest of the theme in the application. The progress bar itself is not a motivational factor; however, as progress and clear visual feedback both are motivational factors [9, 11], it can be used to visualize progress for the users and thus motivate them to proceed through the entire application and all of the scenarios.
7. Visual theme resembles the topic. The visual theme of the application was determined to match the topic; therefore, red and white were used for most of the text and buttons in the game. These colors are generally associated with first aid (see

Fig. 1 as an example). The background color was decided to be a neutral blue color to avoid a conflict with the red buttons and make interactive objects more noticeable.

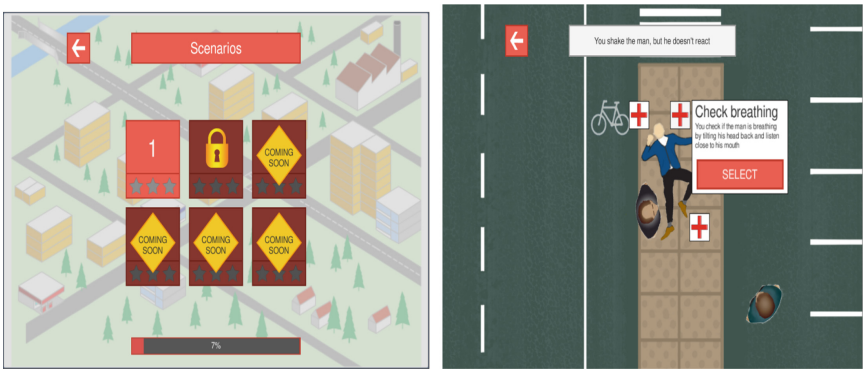

Fig. 1. Two screen shots from the game. Accomplishments and progress bar (*left*) and one of the scenarios with implemented questions and options (*right*). (Color figure online)

3 Methods

Unity3D was used to develop the application. Two different versions of the application were used for testing: one with gamified elements (gamification) and one without gamification elements. The gamified version was implemented with challenges, achievements, and rewards (as described in Sect. 2). A between-group test design was used in which each participant tried either of the two versions. The participants were not told to which group they belonged or what the test was evaluating. In total, there were 20 participants, with 10 in each group. There was an equal number of male and females in each group.

The application was installed on tablets. As we used camera observations of the participants' facial expressions (described in 3.3), the tablet was inclined by 25° to make more of their faces visible to the camera. It was placed approximately 140 cm in front of the participants and approximately 35 cm below the participant's eye level at an angle of 10°.

Quota sampling was used to recruit participants. The criteria for the target group were set to the following: 1. prior certified knowledge with first aid, 2. aged between 20 and 28, and 3. speaks and understands Danish. The age criterion was derived from the first aid certificate requirements for having a driver's license in Denmark, introduced in 2006. However, the first-aid certificate expires after two years.

All participants were informed about the procedure and the observations and signed an informed consent prior to the test. Participants were secured anonymity with an ID number, and their facial expressions not would be disclosed outside of the researchers

in this study. Participants in the gamified version were labelled with a 'G' in front of the ID number and an 'NG' for the non-gamified group.

3.1 Data Logging

The prototype logged total play time, page visits, progression and stars earned, number of attempts for each scenario, how many times they answered the questions correctly, and incorrect answers for each attempt. The logged total time was compared with the participants' perceived time to detect if they were time distorted. The attempts used for each scenario and the number of correct and incorrect answers were used to see how well participants performed. The progression of stars and number of pages visited indicated how much of the application they explored.

3.2 In-Depth Interviews with Use of a Verbal Numerical Rating Scale

We used individual in-depth interviews to obtain in-depth answers from the participants. The main advantage of the interviews was the flexible design in having the participants verbally elaborate on their answers and that it was possible to ask follow-up questions. During the interview, we used a verbal numerical rating scale (VNRS), which is a projective technique used to have participants express their feelings or thoughts on a specific topic [26]. The VNRS ranged from 0–10; 0 was the lowest score, and 10 was the highest. The participants were asked to verbally rate seven elements from the application with regard to different attributes of engagement [9]: perceived time, aesthetics, feedback, challenge, enjoyment, interest, and goals. They were asked two different VNRS questions within each element. After the interview, participants were given a piece of paper with a scenario similar to those in the application. They were then asked four questions. These questions were asked to state whether they recalled what to do in terms of first aid and CPR.

3.3 Observations

The participants were video recorded while they used the application for detecting and analyzing their here-and-now facial expressions compared to their progression in the application. The purpose was to measure signs of concentration and happiness (and the function of the intended humoristic elements implemented). Happiness can be detected by lip-corner pulls, raising the cheeks, and narrowing the eyes by raising the cheeks [24, 25]. Concentration can be detected by narrowing of the eyes and lowering and bringing together the eyebrows [24, 25].

Krippendorff's Alpha (KALPHA)s for both the interviews and the observations (both with 3 coders) was calculated to check for intercoder reliability [27]. KALPHA for coding the interviews was 0.537 for participants for the gamified version and 0.546 for the non-gamified version. KALPHA for the observations was 0.320 for participants within the gamified version and 0.364 for the non-gamified version.

4 Humor in a Serious Game

Humor in serious games is highly complex, as it can have both different functions and specific experiences and outcomes in game play [20, 21]. However humor is associated with pleasure as an overall function [20] and might cause laughter and moments of shared fun. Besides pleasure, humor might also contain emotional responses, such as playfulness, surprise, or other responses associated with mirth [20]. In the implementation, we wanted to use humor for emotional and cognitive functions [20], both for maintaining engagement as well as attention and awareness. The name of some of the achievements was designed to be humorous. When users reached 50% progress, they were awarded an achievement stating "50% - Whoa, we're half way there," which is a reference to Bon Jovi's song "Livin' on a Prayer." Achievement 6 was unlocked when participants had opened all of the subpages on the "Learn CPR" page and stated "You should consider a job as a lifesaver!" Also, humor elements were implemented in the soundscape, e.g., there was the screeching of tires down the road with the arrival of an ambulance.

The participants that reported the highest amount of perceived time were G3 and NG3. They were also the participants with most positive evaluation in terms of the implemented humor. In the interviews, they stated that the application was "fun and serious but still game-like" (G3) and "had good basics and was more fun than traditional learning of CPR" (NG3). Other participants also mentioned that "it was done in a humorous way" (G2). However, some participants were confused by the humor and did not think the implemented humoristic elements were funny. Participant NG7 stated that she was "confused by the humor which suddenly occurred." This means that there might be a mismatch between the genre expectations and the actual content in the application, but to be humorous, the content must also be a surprise [20]. NG7 was also uncertain whether the purpose of the humoristic elements and scores were partly to compare them with friends (social function) or if the game was actually was an application for testing knowledge level (cognitive function).

Seven participants stated that it was a "fun experience" (G1, G3, G6, NG4, NG5, NG8, and NG10). Furthermore, only two participants did not smile through the usage of the application. However, the observed facial expressions related to participants' happiness did not reveal much difference between the gamified and non-gamified application (Table 1). Based on the observations, it can be questioned how enjoyable the application actually was. The findings of these observations is taken to a methodological level as being very difficult to actually observe and interpret participants' facial expressions. Table 1 shows the coders' ratings of the participants' enjoyment/happiness.

It is rather problematic using only self-reports (e.g., interviews and questionnaires) as the sole method for user feedback and evaluation. Users might have difficulties expressing their feelings and behaviors towards the application, and self-reports are often based on an evaluation after the event or "disturbing" the flow during usage. This is the reason why self-reports with advantages could be supplemented. However, observations and psychophysiological measurements (e.g., electroencephalography and skin conductance) can be very difficult to both set up and interpret in gaming. The idea is to come closer to what users do, feel, and think and guess what they have on their minds.

Table 1. Average number of signs of enjoyment/happiness (three coders), based on FACS codes [24, 25]

	Definition	Gamified	Non-gamified	Difference, N-G
1. Somewhat happy	Small lip-corner pull	2.36	2.17	0.21
2. Moderately happy	Medium lip-corner pull with mouth closed. Raising cheeks	2.2	2.2	0
3. Very happy	High lip-corner pull with mouth open. Narrowing the eyes by raising the cheeks	0.43	0.93	0.50

However, as many scholars have argued, information on causal relationships between discrete brain structures and their putative functions is still rather limited [6, 16].

5 Engaged or Not Engaged

In general, there was not much difference between the gamified and non-gamified group in terms of their VNRS ratings for six of the seven engagement elements (perceived time, aesthetics, feedback, enjoyment, interest, and goal). The only main difference was seen in "challenge," for which the participants were asked how challenging the application was. The question was rated on an average of 5.0 (0 lowest and 10 highest) for the gamified group and 2.9 for the non-gamified group. This could be interpreted as the gamified group feeling more challenged than the non-gamified group. Some participants in the gamified group stated that the achievements and points were a challenge for them to obtain, which might be the reason why they felt more challenged.

We assumed that the participants who used the gamified version would spend more time on the application compared to those who used the non-gamified version. However, the result was opposite. The group that used the gamified version had an average play time of 564 s (perceived time of 654 s) and the non-gamified had an average play time of 619.5 s (perceived time of 648 s). There is no significant difference between the difference in play time and perceived time between the two groups ($p < .2123$, Wilcoxon rank sum test). The minor difference in the two groups' perceived time and play time could be due to the participants not being in the flow [13, 14] or being engaged [9]. One reason could be the intrinsic motivation in the content of the application, but several participants also mentioned external factors, e.g., G7 stated several things that distracted from her experience (e.g., background noises and the artificial set up for the test).

The logged data revealed that the gamified group played scenario 1 more frequently (on average 0.4 more often) compared to the non-gamified group. This could be explained by a higher level of interest within the gamified version for scenario 1 and the fact that scenario 2 was seldom retried by either group. Scenario 2 had too high of a similarity with scenario 1 (not enough new content), which affected the interest. The

average numbers of tries for scenario 2 was 0.7 for the gamified version and 0.8 for the non-gamified version.

The non-gamified group had an average of 78.9% progression, and the gamified group had an average of 77.4%. The gamified group's progression was lower than expected; however, all of the gamified group's participants had visited all of the pages in the application, but two participants of the non-gamified group missed a few. A general comment from the interview was that the questions/challenge were too easy. However, even though that we could not see much difference in the VNRS ratings (except the challenge questions), observations, or logged data, the comments from the interviews highlight major differences between the two groups, e.g., G2 stated, "The progression bar kept me going ...' G1 stated, "I liked and felt motivated by the achievements, and I wanted to unlock all of them ... I wanted to know how I could get all the achievements." However, in both groups, there were explicit statements mentioning that "it was a fun experience" (G1, G2, G3, NG4, NG8, and NG10).

After the interview, all participants were given a scenario on paper regarding a drowning accident to see if they recalled any of the information given in the application (cognitive learning perspective). Of 80 questions (four per participant), only 14 questions were answered incorrectly, and there was no difference between the two groups. The only difference was a total of one more correct answer for the non-gamified group for questions 1 and 4.

6 Conclusion

The topic of first aid had an impact in terms of participants' interest. Several participants stated that they knew it was an important topic but felt they did not need to know anything about it (again) because they believed they would never be in a situation where CPR was needed. This was mostly as expected. The idea behind this study was motivating people already trained in CPR to refresh their first-aid knowledge with the support of a serious game with gamified elements. With regard to this aim, the application was a success. An after-evaluation to test the knowledge gained revealed very few incorrect answers.

However, several topics can be further investigated within both the game design and the methods for evaluating engagement and gamified elements. The scenarios in this game should have been more challenging; nearly everyone who attempted the scenarios scored a high number of correct answers on their first attempt. The lack of challenge in the scenarios affected motivation, interest, and perceived time. If users do not feel challenged, there is less of a chance that they will be engaged in the game and may grow bored and lose interest in the application.

In general, there was not much difference between the gamified and non-gamified groups in terms of both the VNRS ratings from the interviews, time spent, and perceived time from the data logging. However, further research with more participants and better control set up is needed to determine the actual effects of gamified versus non-gamified elements.

Our conclusion reveals that it is rather difficult to both implement and measure humor in a serious game. The elements of humor were evaluated differently, both

among the game developers and test subjects. In that sense, the conclusion is part of the already well-known trap for game developers. Game developers tend to make serious games fun using such heuristics as game interface, game mechanics, game story, and game play [22], but the implemented humor might not match either the overall purpose or the players' interest. However, the triangulated methods for measuring humor provided some insight. The FACS codes used to observe participants' facial expressions combined with the interviews could be one way of measuring humor, but this is rather time-consuming, and training is necessary both for detecting facial expressions and conducting the interviews. Intercoder reliability is an important tool for findings and reporting and in the design process. Besides an indication of reliability, it also gives the design developer a better common understanding of the application, as necessary discussions, definitions, and decisions can be made on the basis of better common understandings due to intercoder indications.

References

1. Wissenberg, M., et al.: Association of national initiatives to improve cardiac arrest management with rates of bystander intervention and patient survival after out-of-hospital cardiac arrest. JAMA **310**(13), 1377–1384 (2013)
2. Committee on the Treatment of Cardiac Arrest: Strategies to Improve Cardiac Arrest Survival: A Time to Act. National Academies Press, Washington, D.C. (2015). Current Status and Future Directions, Board on Health Sciences Policy; Institute of Medicine
3. Zichermann, S.C., Cunningham, C.: Gamification by Design. O'Reilly Media Inc., Sebastopol (2011)
4. Burke, B.: Gamify: How Gamification Motivates People to Do Extraordinary Things. Gartner Inc., Brookline (2014)
5. Schønau-Fog, H., Bjørner, T.: "Sure, I would like to continue": a method for mapping the experience of engagement in video games. Bull. Sci. Technol. Soc. **32**(5), 405–412 (2012)
6. LeDoux, J.E.: A neuroscientist's perspective on debates about nature and emotion. Emot. Rev. **4**(4), 375–379 (2012)
7. Lee, M.: Ticky ends: employing thinly-sliced narratives in serious games for mobile platforms. Int. J. Multimedia Ubiq. Eng. **9**(10), 349–362 (2014)
8. Yannakakis, G.N., Hallam, J.: Real-time game adaptation for optimizing player satisfaction. IEEE Trans. Comput. Intell. AI Games **1**(2), 121–133 (2009)
9. O'Brien, H.L., Toms, E.G.: The development and evaluation of a survey to measure user engagement. J. Am. Soc. Inf. Sci. Technol. **61**(1), 50–69 (2009)
10. Quesenbery, W.: The five dimensions of usability. In: Albers, M.J., Mazur, M.B. (eds.) Content and Complexity: Information Design in Technical Communication. Lawrence Erlbaum, New York (2003)
11. Sweetser, P., Wyeth, P.: Gameflow: a model for evaluating player enjoyment in games. Comput. Entertain. **3**(3), 1–24 (2005)
12. Marczewski, A.: Even Ninja Monkeys Like to Play: Gamification, Game Thinking and Motivational Design. Gamified, Addlestone (2015)
13. Csikszentmihalyi, M.: The Evolving Self A Psychology for the Third Millennium. Perennial, New York (1993)

14. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper Perennial, New York (1990)
15. Keller, J., Bless, H.: Flow and regulatory compatibility: an experimental approach to the flow model of intrinsic motivation. Pers. Soc. Psychol. Bull. **34**(2), 196–209 (2008)
16. Gardhouse, K., Anderson, A.: Objectives and subjective measurement in affective science. In: Armony, J., Vuilleumier, P. (eds.) The Cambridge Handbook of Human Affective Neuroscience, pp. 57–81. Cambridge University Press, Cambridge (2013)
17. Weinberger, M.G., Gulas, C.S.: The impact of humor in advertising a review. J. Advert. **21** (4), 35–59 (1992)
18. Lazzaro, N.: Why we play games: four keys to more emotion without story. In: Game Developers Conference, 8 March (2004)
19. Dillon, R.: On the Way to Fun: An Emotion-based Approach to Successful Game Design. A. K. Peters, Natick (2010)
20. Dormann, C., Biddle, R.: A review of humor for computer games: play, laugh and more. Simul. Gaming **40**(6), 802–824 (2009)
21. Carr, D.: Contexts, gaming pleasures, and gendered preferences. Simul. Gaming **36**, 464–482 (2005)
22. Wang, H., Shen, C., Ritterfeld, U.: Enjoyment of digital games: what makes them "seriously" fun? In: Ritterfeld, U., Cody, M., Vorderer, P. (eds.) Serious Games: Mechanics and Effects. Routledge, New York and London (2009)
23. Ritterfeld, U., Cody, M., Vorderer, P.: Introduction. In: Ritterfeld, U., Cody, M., Vorderer, P. (eds.) Serious Games: Mechanics and Effects. Routledge, New York (2009)
24. Rozin, P., Cohen, A.B.: High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of americans. Emotion **3**(1), 68–75 (2003)
25. Cohn, J.F., Ambadar, Z., Ekman, P.: Observer-based measurement of facial expression with the facial action coding system. In: Cohn, J.A., Allen, J.B. (eds.) The Handbook of Emotion Elicitation and Assessment. Oxford University Press Series in Affective Science, Oxford (2006)
26. Bjørner, T.: Data collection. In: Bjørner, T. (ed.) Qualitative Methods for Consumer Research. Hans Reitzels Forlag, Denmark (2015)
27. Hayes, A.F., Krippendorff, K.: Answering the call for a standard reliability measure for coding data. Commun. Methods Meas. **1**(1), 77–89 (2007)
28. Oak, J.W., Bae, J.H.: Development of smart multiplatform game app using UNITY3D engine for CPR education. IJMUE **9**(7), 263–268 (2014)
29. Kelle, S., Klemke, R., Specht, M.: Effects of game design patterns on basic life support training content. J. Educ. Technol. Soc. **16**(1), 275–285 (2013)