UNIVERSITY OF SKÖVDE

VS-GAMES 2015

7th International Conference on Games and Virtual Worlds for Serious Applications

16–18 September 2015
Skövde, Sweden

IEEE

Catalog Number: CFP1538G-ART (Xplore)
CFP1538G-USB (USB)
ISBN: 978-1-4799-8102-1 (Xplore)
978-1-4799-8101-4 (USB)

2015 2015 IEEE 7th International Conference on Games and Virtual Worlds for Serious Applications (VS-Games)

IEEE Catalog Number: CFP1538G-ART (Xplore)
ISBN 15: 978-1-4799-8102-1 (Xplore)

IEEE Catalog Number: CFP1538G-USB (USB)
ISBN 15: 978-1-4799-8101-4 (USB)

Printed copies of this publication are available from:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com

Picture on the front cover:

The restaurant and view of Skövdes´ Boulogner park in three different techniques.

1910 - Ludvig Ericson. Black-White (large format negative)

2015 – Torbjörn Svensson, Digital color photography (Nikon D700)

2014 – Max Mellhage, computer visualization of Boulogner park (Unreal Engine)
+ hand cursor by Torbjörn Svensson

Montage by Torbjörn Svensson.

Produced by IEEE eXpress Conference Publishing

For information on producing a conference proceedings and receiving an estimate, contact conferencepublishing@ieee.org
http://www.ieee.org/conferencepublishing

A systematic mapping study on gamified software quality

Juan Vargas-Enríquez, Lilia García-Mundo
Institute of Technologies and Information Systems
University of Castilla-La Mancha
Ciudad Real, Spain
{JuanAntonio.Vargas, LiliaCarmen.Garcia}@alu.uclm.es

Marcela Genero, Mario Piattini
Institute of Technologies and Information Systems
University of Castilla-La Mancha
Ciudad Real, Spain
{Marcela.Genero, Mario.Piattini}@uclm.es

***Abstract*— Gamified software is currently very popular, and it is expected that it will be widely adopted over the coming years. The social impact of gamified software will probably be very high, and we therefore believe that the assessment and improvement of gamified software quality may be necessary. The aim of this paper is to present a systematic mapping study (SMS) carried out to discover the current state of the research on software gamification quality, in order to identify gaps that merit rigorous future investigation. This SMS allowed us to select 35 papers found in five digital libraries up to April 2014. This paper summarizes the main issues of the planning and the conducting of the SMS. The main results of the data synthesis are detailed and future work is also outlined.**

Keywords—gamification; gamified software quality; ISO/IEC 25010; systematic mapping study

I. INTRODUCTION

The term "Gamification" originated in the digital media industry in 2008, but was not widely adopted until the second half of 2010. Gamification is defined as the use of game design elements in non-game contexts [1]. Game design elements are constituent parts that are found in most games, readily associated with games, and found to play a significant role in gameplay [1]. Examples of game design elements are: points, levels, leaderboards, badges, challenges/quests and engagement loops, among others [2]. The non-game context refers to the use of elements of games for purposes other than their normal expected use as part of an entertainment game [1]. The non-game context can be as varied as crowdsourcing, social networks, loyalty programs, marketing, industry, education, health initiatives, etc. This definition explicitly excludes serious games, another emerging technology that uses game elements, as serious games are complete games, whereas gamification is a way of designing applications and services [3]. However, the boundary between "game" and "gamification" can often be blurry [1].

Gamification is a valuable approach as regards making non-game products, services or applications more enjoyable, more motivating and/or engaging to use [1]. Gamified Software has undergone a very important growth; Gamified software is currently very popular and this means by which to engage users is becoming a growing trend. It is therefore expected that this key emerging technology will be widely adopted over the coming years. M2 Research estimates that the market spend on gamification solutions, applying game mechanics and behavioral analytics in non-traditional applications will reach 2.8 billion USD by 2016 [4]. According to Gartner's 2013 Hype Cycle Special Report [5], gamification is at the top of the peak of inflated expectations, with 5-10 years of mainstream adoption, which is when "Early publicity produces a number of success stories—often accompanied by scores of failures. Some companies take action; many do not". The growing interest in gamification is also reflected in the academic context [6]. The aforementioned aspects reveal that the social impact of gamified software on users and many areas of society will probably be very high, and we therefore believe that the assessment and improvement of gamified software quality may be necessary. Measures with which to measure gamified software quality characteristics are therefore required. These measures will permit the evaluation of gamified software quality and help determine, for instance, whether one gamification solution for a particular piece of software is better than another.

The goal of the research presented in this paper is to discover the current state of the research on gamified software quality and to identify gaps that merit rigorous future investigation. In order to achieve this goal we carried out a systematic mapping study (SMS) following the guidelines proposed by [7] [8] [9].

The remainder of the paper is organized as follows: Section II presents the related work. Section III provides a description of how the SMS was planned, while Section IV explains how the SMS was conducted. Section V sets out the data synthesis and results. An analysis of the threats to the validity of the SMS is presented in Section VI, and finally, our conclusions and future work are presented in Section VII.

II. RELATED WORK

As related work, we will introduce some existing literature reviews that are to some extent related to that presented herein.

Xu [10] conducted a literature review on web application gamification and analytics. In this review the author surveyed 4 gamified applications and 3 computer games. He concluded that much can be learned from the principles and practices of gaming, although the focus must be kept on the limits and potential traps embedded in gamification.

978-1-4799-8102-1/15/$31.00 ©2015 IEEE

Hamari, Koivisto, and Sarsa [6] conducted a literature review focused on the empirical evidence of the motivational affordances (gamification design elements) implemented and the psychological (enjoyment, engagement, motivation, etc.) and behavioral (participation, effectiveness of learning, etc.) outcomes measured. They found 24 papers published between 2008 and 2013. They concluded that even though gamification seems to have positive effects, these are greatly dependent on both the context in which gamification is being implemented and the users who employ it.

Pedreira, Garcia, Brisaboa, and Piattini [11] conducted a SMS on gamification in software engineering. As a result of the SMS they found 16 primary studies, published between 2011 and 2013. They came to three major conclusions; 1) More research efforts analyzing the impact of gamification on software engineering are needed, 2) Most studies share a lack of methodological support that would make their proposals replicable in other settings, and 3) The integration of gamification into an organization's existing tools is also an important challenge that needs to be addressed in this field.

In a previous work we conducted an SMS focused on serious game quality [12]. In that work we found 112 papers published up April 2013. We came to three conclusions: a) it is necessary to address the quality of serious games from early stages of development, b) more empirical evidence is needed as regards proposals that address the quality of serious games, and c) the experiments need to be replicated by researchers other than those who proposed the serious game.

The literature review presented in this paper is different from those previously described in several respects:

- Its goal: our goal is to collect all the existing literature on gamified software quality, and not only the empirical studies.
- Procedure: the SMS presented in this paper has been carried out using the guidelines proposed for SMSs in [7] [8] [9].
- Period of time: the time period is longer and more recent; more studies have therefore been included: we searched for all papers published up to April 2014 without setting a start date.

III. PLANNING THE SMS

The protocol developed to conduct this SMS includes objectives, research questions, search strategy, selection strategy (inclusion/exclusion criteria), study selection procedure, data extraction strategy, and data synthesis.

The main research question that will lead to the objective of this SMS being achieved is:

What is the state-of-the-art of the research on gamified software quality?

Since the main research question is quite general, it was subsequently divided into 6 research questions (see Table I), in order to provide a more specific understanding of the topic being investigated.

TABLE I. RESEARCH QUESTIONS

Research questions	Main motivation
RQ1. What gamified software quality characteristics have been addressed?	To identify the quality characteristics and to map them onto the quality characteristics proposed in ISO/IEC 25010 [14].
RQ2. What research methods have been used when investigating gamified software quality?	To determine whether or not the research has been validated and to discover what research method was used to validate it.
RQ3. What has the outcome of the research been?	To discover the outputs that are produced when investigating software gamification quality.
RQ4. On which software artifacts from gamified software has the research on quality been focused?	To discover whether gamified software quality has been researched throughout the whole software development lifecycle or whether it has focused solely on certain parts.
RQ5. What gamification elements have been used in gamified software?	To discover what gamification elements have been considered in the initiatives related to gamified software quality.
RQ6. What have the application areas of gamification been?	To determine the application area on which gamified software quality initiatives have been focused.

The search string used to collect the papers from sources was constructed using the steps described in Brereton et al [13]. The major search terms would, at first sight, appear to be “quality” and "gamification", but as quality is a multidimensional term, we decided to use the search term “evaluation” and its related terms, because we found that these terms were frequently used in other papers dealing with software quality. The other major search term was “gamification”, but we also used the terms “gameification”, “ludification” and “funware”, since these terms were also used before the term “gamification” was widely adopted. The major search terms were therefore “evaluation” and “gamification” and we built the search string with the alternative terms within the search terms, or with synonyms, as shown in Table II. We searched in 5 of the most important digital resources: Scopus, Science@Direct, IEEE Digital Library, ACM Digital Library, and Springer. The search string was applied to the title, abstract and keywords.

The papers which were included were those dealing with gamified software quality that had been written in English. We decided to include journals, conferences and workshop papers published up to April 2014 without fixing a starting year in order to make this SMS as complete as possible. We excluded

TABLE II. MAJOR SEARCH TERMS AND THEIR ALTERNATIVE TERMS

Major Terms	Alternative terms
Evaluation	(evaluat* OR assess* OR measur* OR test*)
Gamification	(gamif* OR gameif* OR ludif* OR “funware”)

papers related to serious games or game-based learning, or in which the paper´s contributions was not related to the evaluation/assessment/measuring or testing of quality characteristics.

The mechanics of the entire selection procedure were: the first author reviewed the paper and another author then verified it. Any discrepancies were resolved by means of a consensus being reached between the four authors, taking into account the full text of the paper.

In order to ensure that the same criteria were used for the data extraction and classification of the selected papers, a strategy based on providing a classification scheme based on the research questions was defined. This classification scheme consisted of 6 dimensions (one for each research question) with several dimensions each one. For research question 1, the quality characteristics categories were defined on the basis of the ISO/IEC 25010 standard [14] due to it is the current standard of software product quality. For research question 2, the classification of research methods suggested in [15] was used. The categories for the remaining dimensions were initially defined prior to the data extraction and needed to be refined. A detailed description of the classification scheme can be accessed at http://alarcos.esi.uclm.es/SMS-GamificationQuality/.

IV. Conducting the SMS

Four researchers took part in the whole process. The planning of the SMS began in December 2013. All papers related to gamified software quality published up to April 2014 were retrieved in April 2014. The paper selection process is shown in Fig. 1; 254 papers were found, and of these 39 were discarded because they were duplicate papers (the same paper in a different source). The title and abstract of each paper were subsequently reviewed and the number of papers selected was reduced to 56. Inclusion and exclusion criteria were applied to the full text, and 30 more papers were discarded; 26 papers were eventually selected as primary studies. Some months after our paper selection process finished, the literature review carried out by Hamari, Koivisto, and Sarsa [6] was published. Upon considering the similarity between the topic addressed by [6] and our SMS, we decided to check the papers selected in [6] to discover whether or not they coincided with ours. We found that only 5 of the papers selected in [6] coincided with our selection, and believe that this lack of coincidence may be owing to differences in the search string. However, Hamari, Koivisto, and Sarsa [6] did not explicitly describe the search string. In order to be fair and rigorous, we checked all 19 of the papers that did not coincide and applied the inclusion/exclusion criteria (see Table 4) to them. We found 9 more papers that met the inclusion/exclusion criteria of this SMS, and they were therefore included. Most of the papers from Hamari, Koivisto, and Sarsa [6] that were not included were focused on serious games and not strictly on gamification. A detailed explanation of the reasons that led us to include or exclude the papers in [6] in this SMS, and the list of the 35 primary studies selected in this SMS can be accessed at http://alarcos.esi.uclm.es/SMS-GamificationQuality/.

Although the protocol was initially defined during the planning of the SMS, it was necessary that all the authors review and refine it during the execution of the SMS. The identification and selection of primary studies was performed by the first two authors. In order to reduce the risk of a publication being incorrectly included in or excluded from the SMS, each paper was reviewed by at least two authors. In those cases in which these two authors had conflicting views, it was necessary for a third and a fourth author to review the publication and make a final decision.

Some search engines in databases or digital libraries have limitations when using complex Boolean search strings. When a database or digital library did not allow the use of complex Boolean expressions, the search string was split or modified to accommodate the limitations of the search engine. The aim was to obtain the same results that had been achieved using the original search string.

V. Data Synthesis and Results

In this section, the answers to each of the research questions shown in Table I are presented. Some other additional results are detailed at the end of the section.

A. *RQ1. What gamified software quality characteristics have been addressed?*

The process used to match the characteristics shown in the ISO/IEC 25010 standard [14] with the characteristics investigated in each paper is described as follows. The full text of the paper was read in order to search for the quality characteristics that were addressed by the researchers, and the standard was then viewed in order to find the characteristic or characteristics that, in the author's opinion, best matched the characteristics found in the paper. In the review of the full text of the selected papers it was found that the majority of them did not explicitly mention that an assessment of the quality of gamification was being made. The results revealed that most researchers are interested in assessing various aspects of the use of the gamified software, and particularly its effectiveness as regards motivating and engaging users in desired behaviors.

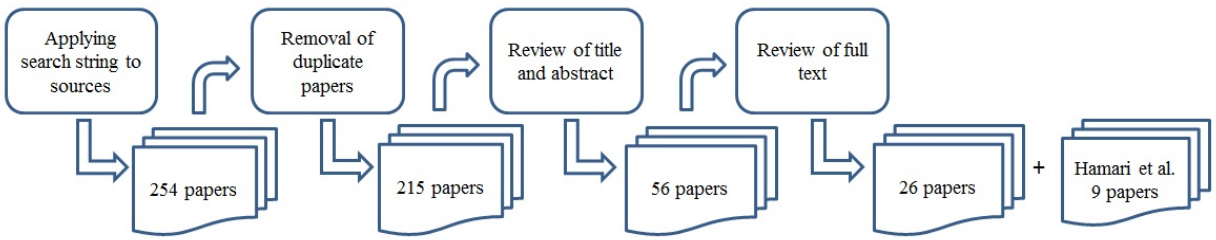

Fig. 1. Paper selection process

Another of the most frequently evaluated aspects was user satisfaction with the gamified software. Most of the papers (32, 91.43%) dealt with the quality characteristics of the quality in use model. The characteristic most frequently addressed was effectiveness (30 papers, 85.71%). In these papers, effectiveness was a measure of how the gamified application achieved its goals to motivate and engage users to achieve the desired behavioral change. These results were then compared with those concerning non-gamified software to observe whether there was any kind of improvement. Another important finding is that the main factor behind the use of gamification is user engagement (12 papers, 34.29%), followed by engagement/motivation (9 papers, 25.71%) and motivation (9 papers, 25.71%). The remaining 5 papers did not mentioned why gamification was used. In papers focused on engagement, the researchers attempted to encourage users to continue doing certain tasks or activities, while in those focused on motivation the researchers wished to increase the users' participation in certain tasks or activities. Following effectiveness is satisfaction (20 papers, 57.14%). The most frequently investigated sub-characteristics as regards satisfaction were pleasure (18 papers, 51.43%) and usefulness (8 papers, 22.86%). These papers assess the effect of gamification on several aspects of user satisfaction such as enjoyment, fun, or the perception of usefulness of the gamified software. A summary of the quantitative results of the characteristics of the quality in use model is shown in Table III.

Only 11 papers (31.43%) dealt with the quality characteristics of the product quality model. The characteristics of the product quality model that were most frequently researched were usability (8 papers, 22.86%) followed by functional suitability (3 papers, 8.57%). The most frequently investigated sub-characteristic of usability was operability (7 papers, 20%), followed to a far lesser extent by the sub-characteristics of user interface aesthetics and learnability (1 paper each). The papers related to usability, focused on the usability of the entire application, not just those aspects related to gamification. A summary of the quantitative results of the characteristics of the product quality model is shown in Table IV.

A detailed description of the process used to match the characteristics in the ISO/IEC 25010 standard [14] with the characteristics investigated in the papers can be accessed at http://alarcos.esi.uclm.es/SMS-GamificationQuality/.

TABLE III. DISTRIBUTION OF PAPERS ACCORDING TO CHARACTERISTICS OF THE QUALITY IN USE MODE

Quality Characteristic	Number of papers	Primary studies references
Effectiveness	30	P1, P2, P4, P5, P6, P7, P8, P9, P10, P11, P12, P13, P15, P16, P17, P18, P19, P20, P21, P23, P24, P25, P26, P27, P28, P29, P30, P31, P33, P34
Satisfaction-Pleasure	18	P1, P2, P4, P5, P6, P7, P8, P12, P13, P15, P17, P20, P26, P28, P29, P31, P32, P35
Satisfaction-Usefulness	8	P5, P7, P17, P23, P25, P28, P31, P32
Total	**32**	

TABLE IV. DISTRIBUTION OF PAPERS ACCORDING TO CHARACTERISTICS OF THE PRODUCT QUALITY MODEL

Quality Characteristic	Number of papers	Primary studies references
Functional suitability-Functional appropriateness	3	P3, P22, P29
Usability-User Interface Aesthetics	1	P17
Usability-Operability	7	P5, P7, P12, P13, P14, P31, P35
Usability-Learnability	1	P14
Total	**11**	

B. RQ2. What research methods have been used when investigating gamified software quality?

The classification scheme of research methods proposed by [15] was used as recommended in [9]. This classification scheme makes it possible to classify empirical research into either validation or evaluation, and non-empirical research in the categories of proposal papers, philosophical papers, opinion papers and personal experience papers. The definitions of these categories can be found in the classification that can be accessed at http://alarcos.esi.uclm.es/SMS-GamificationQuality/.

The results of RQ2 showed that most of the papers are empirical studies (32 papers, 91.43%), and of these, 28 (80%) fall into the category of validation and 4 (11.43%) into the category of evaluation (see Fig. 2). In validation papers the research methods used were quasi-experiments (18 papers, 51.43%) or experiments (10 papers, 28.57%). Most of the quasi-experimental papers used a between subject design with only a posttest (P9, P13, P23, P24, P28, P29), while the remainder used a between subject design with pre and posttests. In all those papers in which an experiment was

TABLE V. CLASSIFICATION OF THE QUANTITATIVE MEASURES OF ENGAGEMENT

Categories	Primary studies references	Measure example
Participation	P2, P5, P7, P11, P19, P23, P24, P29, P34	- number of interactions with the learning platform
Performance	P2, P4, P5, P6, P7, P8, P9, P10, P12, P13, P15, P16, P17, P18, P19, P20, P21, P23, P24, P25, P26, P27, P28, P29, P30, P31	- the percentage of branch coverage of test cases
Communication /socialization	P2, P11, P16, P21, P31, P34	- contributions to forums and other participative media
Frequency of achievements	P6, P16, P17, P18	- Badges awarded

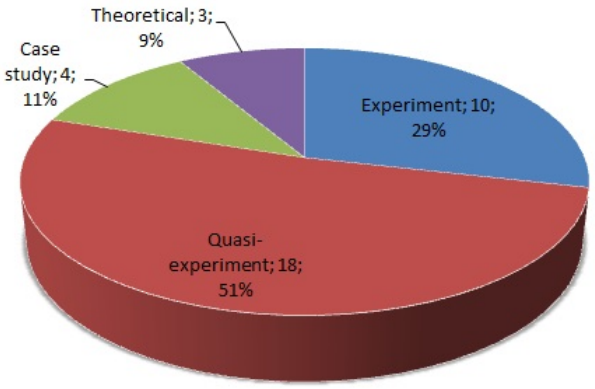

Fig. 2. Distribution of papers by research method

conducted, the researchers used a between subject design with a posttest. In this type of design, the participants in the treatment group used the gamified software and the participants in the control group used the non-gamified software, after which either they were required to fill out a posttest, or usage data were extracted and analyzed. Experimental and quasi-experimental designs were used to assess the effectiveness of the gamified approach, mainly using various quantitative measures applied to the usage data. The objective of these measures was to measure several aspects of user engagement. Although these measures vary according to the type of application, they could be classified into the categories proposed in Table V. The measures of participation aim to assess the user's involvement with the gamified application or the effect of the gamified application on users in terms of the number of actions or transactions performed, while measures of performance are intended to assess how well certain tasks or desired behavior are fulfilled. Communication/socialization measures are intended to assess the exchange or sharing of comments, opinions, or information among users inside or outside the gamified software, while the frequency of achievement measures aim to directly measure engagement through the quantification of points, badges and generally all kinds of achievements.

Satisfaction, usability and functional suitability were in most cases assessed using surveys based on Likert scales.

The evaluation papers carried out case studies (4 papers, 11.43%). A summary of the empirical studies found in this SMS can be accessed at http://alarcos.esi.uclm.es/SMS-GamificationQuality/. With regard to the non-empirical papers, all of them (3 papers, 8.57%) are philosophical papers that propose frameworks for the design and evaluation of gamified applications (P1, P3) or for the design and evaluation of gamification elements (P22).

One positive aspect found is that nearly all of the primary studies (32 papers, 91.43%) are empirical; in all of them, however, the empirical studies were conducted by the same researchers who had proposed the gamification, and none of the studies have been replicated.

C. *RQ3. What has the outcome of the research been?*

The results showed that there are only two types of research outcomes in the papers reviewed. The most common outcome is "knowledge" (30 papers, all papers excluding P1, P3, P13, P22 and P33, or 85.71%), followed to a far lesser extent by "framework" (5 papers, 14.29%). Papers whose outcome is "knowledge" do not present something whose result is "tangible" but rather use empirical studies to confirm whether gamification works, and the outcome of the research in these papers is the "knowledge" about the evidence acquired. With regard to the "framework", 3 papers are presented for the design of gamified applications (P1, P3, P13), one for the design of gamification elements (P22), and one (P33) is a conceptual framework that can be used to explore purchase intentions.

D. *RQ4. On which software artifacts from gamified software has the research on quality been focused?*

The results showed that 32 (all papers excluding P3 and P22, or 91.43%) papers dealt with gamified software quality after the product had been developed, or when a final version was ready. Only 2 (7.69%) papers dealt with the gamified software quality at the design stage. One paper presents a framework based on Bartle's [1] model of player types and on Maslow's [17] hierarchy of needs, and is intended to evaluate the engagement potential of a game design by mapping player motivations onto the various game mechanics (P3), while the other is a framework consisting of a definition of "game achievement" and a three-part model for the evaluation and design of game achievements (P22). As additional results we found that 57.14% of the implementations of the gamified approach are carried out along with the development of the application whilst in 42.86% of the cases this is done by means of a plug-in. The results show that the evaluation of quality in current gamification development practices is often put off until the later stages of the life cycle. These results are not surprising, since researchers have focused on determining the effectiveness of gamified applications and this usually occurs (as is expected in a quality-in-use evaluation) once the software product has been completed or is at least a fully functional prototype. Another possible reason why software gamification quality has focused on the final product is that nearly 43% of gamification implementations are carried out using pre-coded plug-ins. Nonetheless, we consider that it is desirable to consider quality from the early stages of the development of the gamified applications in order to obtain higher quality gamified applications.

E. *RQ5 What gamification elements have been used in gamified software?*

We found 15 different types of gamification elements in the primary studies reviewed. The gamification elements most frequently used are points and badges (23 papers each, 65.71%), closely followed by leaderboards (18, 51.43%) and levels (14, 27.45%), and to a much lesser extent rewards (6 papers, 17.14%) and unlocking (2 papers, 5.71%). There are other gamification elements, but these were mentioned in only one paper each and were classified in the 'other' category (see Fig, 3). Several papers specifically addressed the use of badges, 4 investigated their effects and 1 their design; 3 papers focused

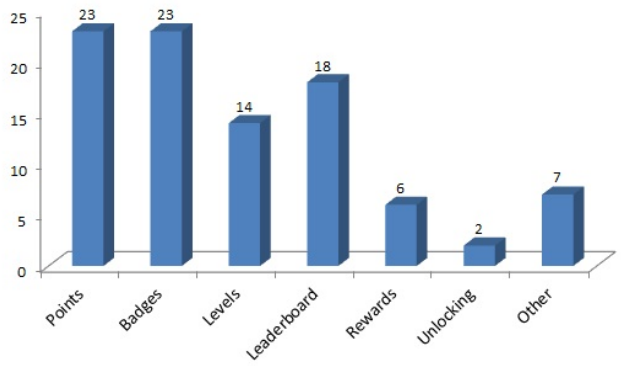

Fig. 3. Distribution of papers by use of gamification elements

on investigating the modification of user behavior (P16, P17, P18), and one paper investigated the effect of badges on user engagement (P6) while another focused on the design of badges (P22).

The results also showed that the combination of points and badges is that most frequently used (16 papers; 45.71%); another frequently used combination is that of points, badges and levels (8 papers; 22.86%). With regard to the use of the gamification elements, several papers claim to deal with gamified software. However, they have only used one or two gamification elements, which are mainly badges or points. We believe that these applications could be improved by adding other gamification elements that could encourage desirable behaviors, thereby making the application more effective [1].

Although it would be interesting to know the level of gamification of the implementations reviewed in this study, unfortunately this is outside of the scope of this paper, owing to the fact that the majority of the papers barely mention how these applications were designed and developed. The researchers have mainly focused on the use of gamified software rather than on its design or development, as is shown by the results of research question 4.

F. RQ6. What have the application areas of gamification been?

The results showed that the main application area of gamification is "Education" (13 papers, 37,14%), and most of these papers deal with gamified courses (11 papers), that is, the use of gamification elements such as points, badges and leaderboards, amongst others, to motivate and engage students in activities such as attending lectures, workshops or labs, assignment submissions, etc. The other two papers address topics such as orientation for new students (P12) and the development of virtual human patients (P19). The most common subjects of gamified courses are: Programming languages (P17, P18, P29), information and communication technologies (P5, P7), healthcare (P6, P19), software engineering (P24), information systems and computer engineering (P2), mathematics for computer scientists (P15), and computer organization and cloud computing (P23).

The other application areas are, but to a far lesser extent, "Work" with 4 papers (P10, P13, P16, P27, 11.43%), followed by "Environment" with 3 papers (P14, P25, P26, 8.57%),

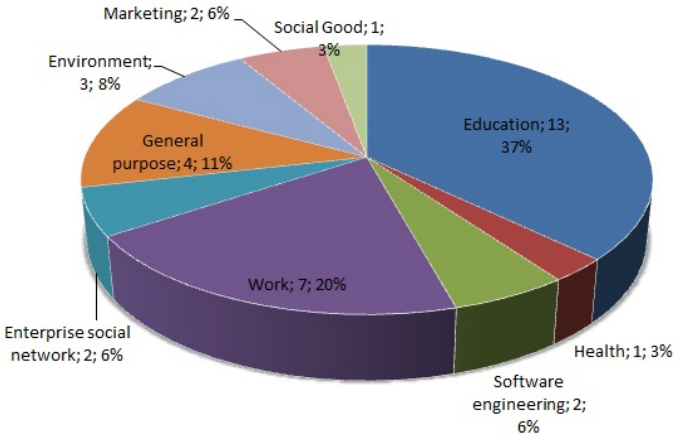

Fig. 4. Application areas

Enterprise Social Network with 2 papers (P11, P34, 5.71%), and Software engineering with 2 papers too (P9, P30, 5.71%). In this fourth category, gamification elements are used throughout the software development life cycle in order to engage and motivate developers when doing their work. The gamified course "software engineering" in the "Education" category should not be confused with the "Software engineering" application area; they are two different things. In the former, gamification elements are used in the teaching process, while in the latter the gamification elements are used during the software development cycle. There are also 4 papers (P1, P3, P22, P33, 11.43%) that do not have a particular application area, signifying that the gamification proposals in these papers are general purpose. It was also found that only 7 papers (P4, P8, P11, P16, P30, P31, P34, 20%) deal with gamification in real world settings, such as social networks, business, enterprises, banking, etc. These results show the importance of gamification and its wide acceptance in the academic world, but they simultaneously highlight the lack of research works on gamified software quality in industry. Figure 4 shows the distribution of papers by application area.

G. Aditional results.

The results reveal that "Education" is the application area in which it is reported that the gamification has had the most

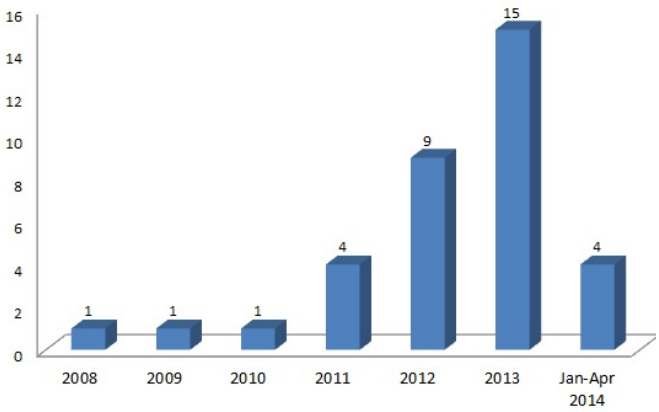

Fig. 5. Distribution of papers by publication year

positive effects, followed by the application area "work". In most of the papers the researchers report the positive effects of gamification in all application areas except the "Enterprise Social Network" category, in which the results were negative owing to the removal of a point based incentive system, and in the "Environment" category in which the results were reported as being inconclusive.

The results also reveal that since 2011, the number of publications related to gamified software has been growing steadily, and doubling every year, as shown in Fig. 5. The number of publications in 2014 was lower, because this study only considered publications up to April 2014. The results therefore show that since 2011, gamified software has undergone a very significant growth, and has in recent years become a "hot topic", thus making software gamification quality an area of opportunity for future research.

VI. Threats to validity

The main threats to validity that should be considered for this SMS are publication selection bias, inaccuracy in data extraction and misclassification [18].

Covering everything that has been written about any topic is impossible. In spite of this, we have done everything possible to gather all the relevant information related to gamified software quality. We therefore also checked the papers included in a related literature review [6] that was published after the execution of the current SMS, resulting in the inclusion of 9 additional papers. With regard to publication selection bias, we searched in 5 of the most important digital resources, including journals, conferences and workshops that are important forums for the disclosure of gamification. We believe that the scope of influence of the journals, conferences and workshops included in this SMS is sufficient to ensure the completeness of the field of software gamification quality. Grey literature such as technical reports or PhD theses were not included since most grey literature originates from or will eventually become peer-reviewed papers. It is possible that some relevant papers have been omitted but this is, to the best of our knowledge, unlikely.

In order to tackle the issue of inaccuracy in data extraction, we defined the research questions in advance, organized the selection of articles as a multistage process, involved four researchers in this process, and documented the reasons for inclusion/exclusion as suggested in [19]. As mentioned previously, the selection of papers to be included as primary studies in this SMS was a rigorous process in which four researchers participated. The data extraction and classification of prose was difficult owing to the lack of a standard terminology for gamification design elements and with which to define quality characteristics in gamification. This may have resulted in some inaccuracies in the data extraction which may in turn have led to a misclassification. However, we believe that the extraction and selection processes were rigorous, as explained in Section IV. We are also of the opinion that the participation of two expert researchers (the last two authors) when performing the classification reduced the risk of misclassification.

VII. Conclusions and future work

The results of the SMS presented herein reveal that research interest in the evaluation of gamified applications has been growing steadily and rapidly since 2011. However, gamification is still an emerging technology and researchers have to date been principally interested in demonstrating its effectiveness, thereby confirming the findings of Hamari, Koivisto, and Sarsa [6]. Although engagement and motivation were assessed in several papers, on most occasions this assessment was not made directly, but by means of an evaluation of the effectiveness of the gamified application as regards achieving its objectives, such as increasing participation or another type of behavioral change.

Although satisfaction and usability have been addressed in several papers, but to a much lesser extent than effectiveness, other quality characteristics have been neglected, especially those related to product quality, such as performance efficiency, compatibility, reliability, etc. Moreover, in most of the papers quality is barely mentioned as a topic of interest for researchers. We believe that this lack of interest in software gamification quality is owing to both the fact that gamification is a new field, and also that existing software quality models are not suitable for the accurate assessment of the particular characteristics of gamified software. We believe it is time for researchers in this area to reach a consensus on which characteristics best characterize the quality of software gamification, and which are the most suitable measures as regards evaluating it. All of the aforementioned aspects highlight the need to adapt existing quality models and thus have a specific quality model with which to assess the quality of gamified software more accurately.

One positive aspect that was found as regards the research method is that nearly all the primary studies (32 papers, 91.43 %) presented experiments. This to some extent reveals that researchers concerned with software gamification initiatives recognize the need to support their proposals with empirical evidence and do not only rely on general wisdom. In most of these papers researchers have reported positive effects of gamification in all areas of application, but in Education is where the best results have been obtained. However, there is a lack of corroboration of the findings obtained by other researchers, signifying that the external replication of the empirical studies is needed in order to obtain more solid and general findings.

It was also clear that very little research has been carried out into the quality of gamified software for businesses or enterprises, and most efforts have focused on the area of education. This shows that there is a gap between industrial practice and academic research that should be addressed by researchers, bearing in mind that gamified software for enterprises represents the biggest segment of the market for new growth, with 25% of the market [4].

The findings obtained have allowed us to identify some possible opportunities for future research which are, among others:

- A quality model for gamified software: Existing software quality models need to be adapted and

extended (if necessary) in order to obtain a consensus on which quality characteristics are relevant as regards evaluation and improvement in the context of gamified software, and how to measure them.

- Need for replication: It is necessary to replicate the empirical studies because in all cases the experimentation was carried out by the same researchers who had proposed the gamified software quality initiative. The material used in the experimentation could therefore be made available in order to encourage other researchers to carry out replications.
- Research in other contexts: Research on gamified software quality should also be carried out in other contexts apart from education, such as businesses and enterprises.
- Address from the early stages: Quality assurance methods that incorporate quality issues from the early stages of the development of the gamified software should be provided.
- Focus on the "good" design of gamified software: Research into what the most effective gamification design elements (or combination of them) are would also be interesting. This task is particularly complex, since the design of gamified software is a multidisciplinary activity that requires skills related not only to software development but also to psychology, among others.

ACKNOWLEDGMENTS

This work has been funded by the following projects: GEODAS-BC (Ministerio de Economía y Competitividad and Fondo Europeo de Desarrollo Regional FEDER, TIN2012-37493-C03-01) and IMPACTUM (Consejería de Educación, Ciencia y Cultura de la Junta de Comunidades de Castilla La Mancha, y Fondo Europeo de Desarrollo Regional FEDER, PEII11-0330-4414). We would also like to thank the Instituto Tecnológico de Ciudad Victoria and PRODEP for granting the two first authors the scholarship that made it possible to complete the research work presented in this paper.

REFERENCES

[1] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining "gamification"," Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, pp. 9-15, September 2011.

[2] G. Zichermann, and C. Cunningham, Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps. O'Reilly, Sebastopol, 2011.

[3] M. Oja, and J. Riekki, Ubiquitous Framework for Creating and Evaluating Persuasive Applications and Games," M. Rautiainen, T. Korhonen, E. Mutafungwa, E. Ovaska, A. Katasonov, A. Evesti, H. Ailisto, A. Quigley, J. Häkkilä, N. Milic-Frayling and J. Riekki (Eds.), Grid and Pervasive Computing Workshops, vol. 7096, pp. 133-140, 2012.

[4] M2 Research, "Gamification in 2012: Market Update, Consumer and Enterprise Market Trends,", 2012. Retrieved September 5th, 2014, from http://gamingbusinessreview.com/wp-content/uploads/2012/05/Gamification-in-2012-M2R3.pdf

[5] Gartner, "Gartner's 2013 Hype Cycle for Emerging Technologies,", 2013. Retrieved September 5th, 2014, from http://www.gartner.com/newsroom/id/2575515

[6] J. Hamari, J. Koivisto and H. Sarsa, "Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification," Proceedings of the 47th Hawaii International Conference on System Sciences, pp. 3025-3034, January 2014.

[7] B. Kitchenham, D. Budgen and P. Brereton, "Using mapping studies as the basis for further research – A participant-observer case study," Information and Software Technology, vol. 53 (6), pp. 638-651, 2011. doi: http://dx.doi.org/10.1016/j.infsof.2010.12.011

[8] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," Technical Report EBSE-2007-01, Software Engineering Group of Keele University Durham UK, 2007.

[9] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic Mapping Studies in Software Engineering," Conference on Evaluation and Assessment in Software Engineering , pp. 26-27, June 2008.

[10] Y. Xu, "Literature Review on Web Application Gamification and Analytics," Technical Report 11-05, Department of Information and Computer Sciences University of Hawaii, Honolulu, HI 2011.

[11] O. Pedreira, F. Garcia, N. Brisaboa, and M. Piattini, "Gamification in software engineering - A systematic mapping,", Information and Software Technology, vol.57, pp. 157-168.

[12] J. Vargas, L. García-Mundo, M. Genero, and M. Piattini, "A Systematic Mapping Study on Serious Game Quality," 18th International Conference on Evaluation and Assessment in Software Engineering (EASE´14), pp. 15, May 2014.

[13] P. Brereton, B. Kitchenham, D. Budgen, M. Turner, and M. Khalil, "Lessons from applying the systematic literature review process within the software engineering domain,", Journal of Systems and Software, vol. 80, pp. 571-583, 2007.

[14] ISO/IEC, ISO/IEC IS 25010: Systems and Software Engineering - Systems and Software Quality Requirements and Evaluation (SQuaRE) - System and Software Quality Models, ISO (International Organization for Standardization), 2011.

[15] R. Wieringa, N. Maiden, N. Mead, and C. Rolland, "Requirements engineering paper classification and evaluation criteria: a proposal and a discussion," Requirements Engineering, vol. 11, pp. 102-107, 2006. doi: 10.1007/s00766-005-0021-6

[16] R. A. Bartle, Designing Virtual Worlds. New Riders, 2004.

[17] A. H. Maslow, Motivation and Personality (Third Ed. ed.), Addison-Wesley, New York, NY, 1987.

[18] D.I. Sjøberg, J. E. Hannay, O. Hansen, V.B. Kampenes, A. Karahasanovic,N.K. Liborg, A.C. and Rekdal, "A survey of controlled experiments in software engineering," IEEE Transactions on Software Engineering , vol. 31 (9), pp. 733-753, 2005. doi: 10.1109/tse.2005.97

[19] J. Liu, J. Dehlinger, and R. Lutz, "Safety analysis of software product lines using state-based modeling," Journal of Systems and Software, vol. 80 (11), pp. 1879-1892, 2007. doi: 10.1016/j.jss.2007.01.047

The Effect of Prior Gaming Experience in Motor Imagery Training for Brain-Computer Interfaces: A Pilot Study

Athanasios Vourvopoulos
Madeira-ITI
Universidade da Madeira (UMa)
Funchal, Portugal
athanasios.vourvopoulos@m-iti.org

Fotis Liarokapis
Human-Computer Interaction Laboratory
Masaryk University
Brno, Czech Republic
liarokap@fi.muni.cz

Mon-Chu Chen
Madeira-ITI
Universidade da Madeira (UMa)
Funchal, Portugal
monchu@andrew.cmu.edu

***Abstract*— Brain–Computer Interfaces (BCIs) are communication systems which translate brain activity into control commands in order to be used by computer systems. In recent years, BCIs had been used as an input method for video games and virtual environments mainly as research prototypes. However, BCI training requires long and repetitive trials resulting in user fatigue and low performance. Past research in BCI was mostly oriented around the signal processing layers neglecting the human aspect in the loop. In this paper, we are focusing at the effect that prior gaming experience has at the brain pattern modulation as an attempt to systematically identify all these elements that contribute to high BCI control. Based on current literature, we argue that experienced gamers could have better performance in BCI training due to enhanced sensorimotor learning derived from gaming. To achieve this a pilot study with 12 participants was conducted, undergoing 3 BCI training sessions, resulting in 36 EEG datasets. Results show that a strong gaming profile not only could possibly enhance the performance in BCI training through Motor-Imagery but it can also increase EEG rhythm activity.**

Keywords—Brain-Computer Interfaces; Serious Games; Virtual Reality; Motor Imagery

I. INTRODUCTION

Brain–Computer Interfaces (BCIs) are communication systems which translate brain activity into control commands [1]. In a BCI system, brain modulation patterns can be extracted and analyzed in order to determine the mental state of the user. These states can be translated with the help of signal processing algorithms and machine learning into a control signal that could act as an input for computers or external devices (e.g. robots). BCI technology is a rapidly growing field of research and has been shown to be very promising for controlling agents within virtual environments [2]–[4].

Currently, three main techniques are used in BCI systems for user interaction and control including: (a) Steady State Visual Evoked Potentials (SSVEP), (b) P300 BCI and (c) Motor-Imagery (MI) or ERS/ERD BCI. SSVEP is caused by visual stimulation by flashing lights to the user and occur at the primary visual cortex of the brain [5]. On the other hand, P300 BCI is generated by measuring the brain evoked response after stimulus onset, positive and negative deflections in the EEG signal after 300ms (hence the name) [6]. Finally, ERS/ERD stands for event related synchronization/desynchronization of the mu (μ) rhythm. Mu is located at the motor and somatosensory cortex of the brain where patterns of electrical activity control voluntary movement [7]. Motor imagery (MI) BCI training is based on visuo-motor imagination and has been widely used as a BCI paradigm in research [8]. MI is relying on the same brain systems that would be used for actual performance of the task by activating the same brain areas as actual action execution or action observation. Results from previous studies have proven mental practice of action to be useful in MI-BCI with beneficial effects in motor control of patients with paralysis [9].

BCI games or neurogames have recently become increasingly more advanced by incorporating immersive virtual environments [10], multiple user objectives, and hybrid control systems integrating both conventional input devices and multiple BCI techniques [11]. Several BCI surveys have analyzed and reviewed BCI games in terms of different approaches, including human–computer interaction (HCI) [12], and immersive virtual reality (VR) [13]. Unfortunately, the fundamental issue of BCI illiteracy/skill where, regardless of the duration of the training session, users are unable to have a stable control, is present in MI-BCI.

The aim of this paper is to examine the effect that prior gaming experience has at the brain pattern modulation during MI training in order to identify the elements that contribute to high BCI control. The hypothesis is based on that experienced gamers could have better performance in MI-BCI training due to enhanced sensorimotor learning derived from gaming. A pilot study with 12 participants, undergoing 3 MI-BCI training sessions was performed. Initial results indicate that a strong gaming profile could possibly enhance the performance in classification accuracy during a MI-BCI training, and additionally increase EEG rhythm modulation in the Alpha, Beta, Theta and Gamma bands. In addition, a relationship with the demographic data provides useful pointers for the trainee profile and its effect to the training outcome.

The rest of the paper is structured as follows. Section II presents relevant background information with several case studies. Section III describes the research hypothesis whereas

978-1-4799-8102-1/15/$31.00 ©2015 IEEE