

A Systematic Mapping Study on Serious Game Quality

Juan A. Vargas¹, Lilia García-Mundo¹, Marcela Genero², Mario Piattini²

¹Instituto Tecnológico de Cd. Victoria, Cd. Victoria, Tam., México

{JuanAntonio.Vargas, LiliaCarmen.Garcia}@alu.uclm.es

²Institute of Technologies and Information Systems, University of Castilla-La Mancha
Ciudad Real, Spain

{Marcela.Genero, Mario Piattini}@uclm.es

Abstract. This paper present a summary of the paper published in the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE 2014). The aim of this research is to discover the current state of Serious Games (SGs) quality initiatives, identifying gaps that merit future investigation. For this purpose, we conducted a systematic mapping study (SMS) on SG quality. The main results are summarized in this paper.

Keywords: Serious Games, Quality, ISO 25010, Systematic Mapping Study.

1 Introduction

Although research has been conducted on several topics related to SGs, more extensive research is needed on SG quality [1,2,3,4,5,6]. SGs are critically important at present, as they can be a means to achieve relevant goals, from both a personal and an institutional point of view. They may be used in fields as diverse as defense, education, scientific exploration, health care, emergency management, city planning, engineering, religion, and politics. In addition, the number of users of these systems grows each day, signifying that their social impact is very high. It is for this reason that SG quality is so critical; they are not just another variety of software (in which it is already assumed that quality is important). They can have a major impact on many areas of society and on many users, and it is therefore our duty as researchers and computer professionals to ensure their quality.

In order to discover the current state of SG quality initiatives and to identify gaps that merit future rigorous investigation, we decided to conduct a systematic mapping study (SMS) following the guidelines proposed by Kitchenham and Charters [7]. The research questions formulated in this SMS are shown in Table 1.

Table 1. Research questions

Research questions	Main motivation
RQ1. What particular quality characteristics of SGs have been investigated by researchers?	To identify the quality characteristics of SGs that have been addressed by researchers, and map them onto the quality characteristics proposed in

	the ISO/IEC 25010 [8].
RQ2. What is the nature of the research on SG quality?	To discover what proposals have been produced by the research work on SG quality.
RQ3. What research methods have been used to investigate SG quality?	To determine whether or not the research has been validated and to discover which research method was used to validate it.
RQ4. What software artifacts has the research on SG quality been focused on?	To discover whether SG quality has been researched throughout the whole software development lifecycle or whether it has focused solely on certain software artifacts.
RQ5. What have the application areas of SGs in the research of SG quality been?	To identify the SG application areas in which quality has been investigated.

The main goal of this paper is to summarize the main results obtained in this SMS and outline future work. The remainder of the paper is organized as follows: Section 2 sets out the results, and in Section 3 the conclusions are presented, along with our future work.

2 Results

After thoroughly applying the guidelines proposed in [7], we finally selected 112 papers, which are listed in <http://alarcos.esi.uclm.es/SMS-SeriousGamesQuality/>

In this section we present the answers to each of the formulated research questions.

RQ1. What particular quality characteristics of SGs have been investigated by researchers?

We match the quality characteristics investigated with the characteristics of the ISO/IEC 25010 standard [8]. The table of 112 primary studies which contains this correspondence is available at <http://alarcos.esi.uclm.es/SMS-SeriousGamesQuality/>.

The results for RQ1 revealed that most of the papers selected addressed more than one quality characteristic or sub-characteristic. We found that the quality model most frequently investigated, in 88 papers, is the quality in use model, as shown in Fig. 1. We also found that 43 of the articles researched a particular characteristic or sub-characteristic of the product quality model (Fig. 2).

The characteristics most frequently addressed by the quality in use model were effectiveness (78.57%) and satisfaction (64.29%). Satisfaction, was mostly addressed by the sub characteristic pleasure (62.50%), and to a far lesser extent by the sub-characteristic utility (13.39%) (Fig. 1). The characteristics of the quality product model most frequently researched were, meanwhile, usability (45.54%) distantly followed by functional suitability (8.93%). We observed that usability was most fre-

quently researched through the use of the operability sub-characteristic (38.39%), closely followed by the user interface aesthetics sub-characteristic (35.71%), and to much lesser degree by the learnability sub-characteristic (8.93%) (see Fig. 2). These results could be explained by the fact that researchers are principally concerned with demonstrating or confirming that the SG meets the (serious) purpose for which it was created, which is why they research effectiveness and usefulness. But researchers are also interested in knowing whether the SG is capable of providing enjoyment and entertainment, which is the part of the SG as regards playability. That is why pleasure, the user interface aesthetics and operability are also the focus of research. We believe that this explains why the other characteristics of SG quality that are not related to these aspects, such as efficiency, performance efficiency or security, have been neglected. Similar findings were reported by Connolly et al [1], who found that the most frequently researched issue was the effectiveness of knowledge acquisition, but that many papers also reported enjoyment and engagement. We also found that very little research has been conducted on the relationship between the effectiveness of SGs and the characteristics of playability that make them effective. It seems that researchers are very interested in knowing whether SGs are effective but not in what makes them effective.

We believe that it would be interesting to investigate which playability aspects have an influence on SGs' effectiveness. Wouters et al [6] came to similar conclusions when stating that a better understanding of the underlying motivational processes such as enjoyment and engagement in SGs is required.

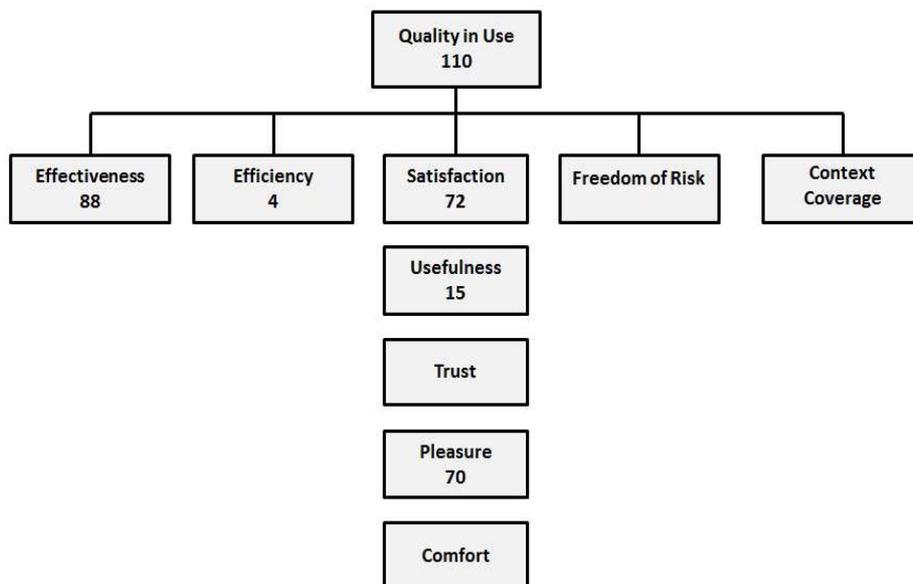


Fig. 1. Distribution of papers according to characteristics of the quality in use model

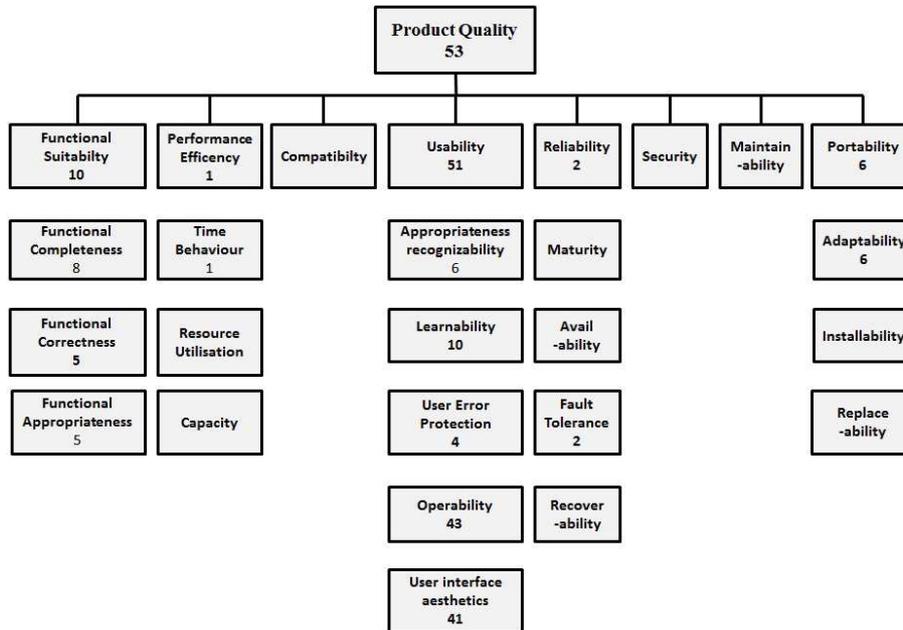


Fig. 2. Distribution of papers according to characteristics of the product quality model

A summary of the quantitative results of the research questions from RQ2 to RQ5 is presented in Table 2.

Table 2. Summary of the quantitative results from RQ2 to RQ5

Research question	Possible answers	Papers	Percentage
RQ2. Research result	questionnaire	43	38.39
	knowledge	37	33.04
	scale	3	2.68
	guide	2	1.79
	tool	2	1.79
	heuristic	6	5.36
	framework	13	11.60
	method	6	5.36
RQ3. Research method	Empirical evidence:	41	36.61
	Validation	39	34.82
	experiment	26	23.21
	quasi-experiment	13	11.61
	evaluation	2	1.79
	Non Empirical evidence:	71	63.39
	philosophical	12	10.71
	proposal	59	52.68

RQ4. Software artifact	requirement	2	1.79
	design	8	7.14
	code	2	1.79
	product	109	97.32
RQ5. Application area	education	68	60.71
	skills development	8	7.14
	awareness	1	0.89
	health	8	7.14
	training	13	11.62
	general application	4	12.50

3 Conclusions and Future Work

In this paper we have presented an SMS related to SG quality; we selected 112 papers from the 1236 found in 6 digital libraries until April 2013. Results also show that SG quality has undergone a very important growth, rising from 3 papers in 2007 to 34 papers in 2012. In recent years SG quality has therefore become a “hot topic”, thus making SG quality an area of opportunity for future research.

The results show that researchers are mainly concerned with demonstrating or confirming the effectiveness of SGs in addition to their capability of providing enjoyment and entertainment, but that very little research has been carried out as regards the characteristics of playability that make SGs more effective. Other characteristics of SGs have barely been addressed, such as efficiency, performance efficiency or security. Since effectiveness and playability are evaluated in the final product there is a need to provide quality assurance methods that incorporate quality issues from the early stages of the SG development focusing, for example, on quality characteristics that may have an impact on an SG’s effectiveness, such as pedagogical and playability aspects introduced in the design stage. Approximately half of the proposals that deal with SGs quality have been empirically validated by means of experiments carried out by the same researchers who propose the SG, and they have not been replicated. Although 28.57% (32 papers) of the studies produced have an outcome that can be applied to any SG, only 2.68% (3 papers) of these outcomes have been validated.

Our interpretation of the review results has allowed us to identify some possible research opportunities:

- Propose, apply and validate a quality assurance method that incorporates quality evaluation from the early stages of the SG development cycle. It would preferably be possible to apply this method to any kind of SG.
- Investigate which playability aspects have an influence on SGs’ effectiveness.
- Carry out more empirical validation. Both internal and external replications are needed in order to corroborate and generalize the findings obtained.

We also plan to carry out a Systematic Literature Review in order to synthesize the empirical evidence on SG quality.

Acknowledgments

This research has been funded by the GEODAS-BC project (Ministerio de Economía y Competitividad and Fondo Europeo de Desarrollo Regional FEDER, TIN2012-37493-C03-01). We would also like to thank the Instituto Tecnológico de Ciudad Victoria and PROMEP for granting us the scholarship that made it possible to complete the research work presented in this paper.

References

1. Connolly, T., Boyle, E., MacArthur, E., Hainey, T., and Boyle, J.: A Systematic Literature Review of Empirical Evidence on Computer Games and Serious Games. *Computers and Education*, 59, 661-686 (2012).
2. Hays, R.T.: The Effectiveness of Instructional Games: A Literature Review and Discussion. Technical Report 2005-004. Naval Air Warfare Center Training Systems Division (2005).
3. Kirriemuir, J., and McFarlane, A.: Literature Review in Games and Learning. Report 8. Graduate School of Education of University of Bristol (2004).
4. McClarty, K.L., Orr, A., Frey, P.M., Dolan, R., Vassileva, V., and McVa, A.: A Literature Review of Gaming in Education. Research Report. Pearson Education (2012).
5. Susi, T., Johannesson, M. and Backlund, P.: Serious Games – An Overview. Technical Report HS- IKI -TR-07-001. School of Humanities and Informatics University of Skövde Sweden (2007).
6. Wouters, P., Van der Spek, E., and Van Oostendorp, H.: Current Practices in Serious Game Research: A Review from a Learning Outcomes Perspective. In Connolly, T., Stansfield, M., and Boyle, L. (eds.) *Games-Based Learning Advancements for Multi-Sensory Human Computer Interfaces: Techniques and Effective Practices*. : Information Science Reference. doi:10.4018/978-1-60566-360-9.ch014, pp. 232-250, IGI Global (2009).
7. Kitchenham, B., and Charters, S.: Guidelines for performing Systematic Literature Reviews in Software Engineering. Technical Report EBSE-2007-01. Software Engineering Group of Keele University Durham UK (2007).
8. ISO.: ISO/IEC FDIS 25010: Systems and Software Engineering - Systems and Software Quality Requirements and Evaluation (SQuaRE) - System and Software Quality Models. ISO (International Organization for Standardization) (2010).