

Comparative Study to Develop a Tool for the Quality Assessment of Serious Games Intended to be used in Education

<http://dx.doi.org/10.3991/ijet.v9i9.4150>

Yassine EL Borji and Mohamed Khaldi
Abdelmalek Essaadi University, Tetouan, Morocco

Abstract—Serious games are present today in several fields such as education, government, health, defense, industry, civil security and science. In our research we are particularly interested in serious games used in education since they have become an essential part of the learning process and one of the main pedagogical tools used for diffusion of educational messages. So the use of serious games in an educational context can motivate students and stimulate their interest as they fit into a world they know and in which they feel comfortable. But the real issue is that they have not been exploited to their full potential in order to ensure efficient use in education.

This work tries to meet this issue. It is interested in contributing to the knowledge of serious games as well as to promote their use in educational field. Based on the metadata schema «SG-LOM» to formally describe serious games and other comparative study, we suggest examining pedagogical, Playful and Technical criteria to conduct this study which aims to develop an assessment and analysis grid of the quality of serious games intended to be used in educational field.

Index Terms—Assessment Grid, Education, Serious Games, SG-LOM

I. INTRODUCTION

In recent years serious games are playing an increasingly important role in the educational sphere, they have reached a certain maturity to become a possible alternative to traditional methods of learning [2]. The integration of information and communication technologies in education through hardware, software, Internet connections, digital content and training for teachers has not proved sufficient to disrupt teaching methods. So we try to meet this need for innovation and motivation through the adoption of serious games in education since we know the vocation of a serious game to make learning attractive by introducing scenarios and playful interactions also young people today are familiar with new technologies and virtual worlds, a habit that allow these young people to be immediately in the heart of the matter since they have most often a video game culture as well as the keys to the gameplay culture which facilitates therefore the transmission of the main educational messages.[1]

It would however be unrealistic to assume that serious games can provide an immediate and perfect solution to the current problems of innovation and student interest, although there are many advocates of the contribution and potential of serious games as the role they could play in formal education. The main challenge we face is to exam-

ine the educational, technical and playful aspects of serious games to determine the criteria for the selection and use in education.

Through this paper, we propose to make a contribution in this area based on different evaluation studies and classification of serious games as well as the metadata schema "SG-LOM"[1] which provide an explicit metadata with new specific fields for serious games by developing an assessment and analysis grid of the quality of serious games intended to be used in education field which allows us to identify the added value of every serious game by highlighting the balance between learning and playful aspects without ignoring their technical side. So the purpose of this grid is to contribute to the knowledge of serious games to promote their use in education.

II. SCIENTIFIC CONTEXT AND OBJECTIVE

For a long time, educators tended to ignore games as a source of education, today the rapid development of games industry, their immense effect on the new generation and the existence of certain characteristics within them possibly advantageous for educational purposes, have led to an increase in the interest shown by the educational community, principally in the last decade, in their use in the field of education.[2]

However, serious games for education are also characterized by their variety in terms of both content and quality, so the evaluation and examination of their educational, technical and playful aspects is a prerequisite before any introduction in education. It will also be necessary for the educators to determine whether or learning methods related games are correlated with aspects teaching of the corresponding instruction.

In his thesis Julian Alvarez [3] also sought to define the basics of a good serious game in education. Thus, according to him, it must:

- Be a coherent system both in substance and form.
- Be adapted to the context of use.
- Have good ergonomic criteria.
- To convey the best fun and educational dimensions.
- Have a good lesson plan.
- Provide an editor to be adapted by the teacher.

According to Julian Alvarez, a very good serious game should be able to adapt their patterns to each learner profile, both with regard to their learning level than its playful video culture, and therefore its ability to use it.

In a previous work [1] which aims also to contribute to the knowledge of serious games and to promote their use in education, we introduced a new metadata schema to describe serious game as learning resources. We proposed a new application profile (SG-LOM) specified for serious games based on the LOM standard in which we adopted a methodological approach that focuses on the study of the main components of serious games taking into account the classification studies and evaluation have been done before.

In this context and taking into account the basics already stressed our objective is to develop a tool intended to assist those in the field of education to assess the (pedagogical, technical and playful) quality of serious games intended to be used in education.

III. SECTIONS OF THE PROPOSED GRID

The grid is divided into four sections each section is composed of elements "subsections" every one of these elements is associated with a set of criteria [4] in order to identify the different aspects constituting a serious game and describes each separately in its own context: educational, technical and playful aspect.

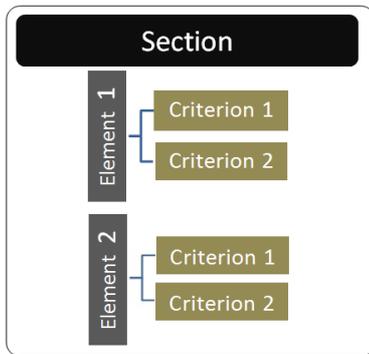


Figure 1. Descriptive diagram of the assessment grid

We summarize the purpose and content of each section:

A. Section I : Serious Game Identification

The objective of this section is to assist in identifying the serious game; it offers a global vision of the game. This section covers the general features of the game [1] such as the name with which the game is known its developer whatever a team or company as well as the type (it's about the game content: Drama, Horror, fantasy, mystery, science fiction ...) and genre (specifies how the game plays: strategy, action, adventure ...) of the game. In the same section other features are required like age group or ESRB rating to ensure clear labeling of contents games for the age group to which they are best suited and game experience required which gives an idea on the level of experience required to play this game as well as the game license and its language besides keywords that characterizes the game.

B. Section II: Pedagogical Specifications

Three major elements form this section: Content, Strategies and Assessment Method, it is the heart of the tool. At first, the information gathered through this section helps people in education to situate the game from an educational and pedagogical point of view. In a second step, they facilitate the identification and adaptation of content in educational context.

1) Content

The content element that contains the following Criteria: Accuracy of the serious content, Consistent with the target audience, Degree of difficulty of activities and Feedback assesses whether the structure of the content of serious game promote its use in a pedagogical context through the presence of consistency between the objectives and content of the game, appropriate structuring knowledge, logic relevance of content based on audiences, presence of various issues and situations with different degree and of complexity and an effective feedback mechanism in time related to each learning task. [6]

2) Strategies

We evaluate through this element the effectiveness of teaching strategies adopted in the game. It covers a set of criteria such as Specification of objectives, Existing Knowledge and representations, Pedagogical Approach and Ability to adapt to each learner profile.

To verify whether the strategies adopted meets the educational needs of the game [5], first they must perceive that all learners have different needs and strengths in a second step if they offer activities that stimulate their commitment and reinforces their learning based on techniques, methods, approaches and different teaching styles to manage different learning styles taking into account every learner profiles.

3) Assessment Method

Here we are interested in one of the major elements that form this analysis grid. The Assessment Method element is interested in the process of collecting and analyzing information from multiple and diverse sources [8] as well as the progression of learners/players and evaluation of their skills to develop a deep understanding of what students know, understand and can do with their knowledge as result of their educational experiences.

This element covers a set of criteria such as Assessment procedure, Progress indicator, Player Tracking, Presence of tutor and Assessment Type. Each of these criteria is responsible for the evaluation of a specific side of the assessment mechanisms in serious games.

C. Section III : Playful Specifications

The third section discusses the playful dimension of the game, it's formed by two major elements: Attractiveness (Fun) and Playability Aspects. It provides a detailed description of the factors of motivation and interaction induced by the game.[9]

The objective of this section is to check whether the graphic areas and scenarios of the game are exciting and attractive to keep the attention of the player manually and intellectually active.

1) Attractiveness (Fun)

The playful side represents a very important part in serious games to ensure the motivation of the player / learner as well as his immersion in the game. So assessing the attractiveness of the game is an obligation to check if there is a balance between playful dose in the game compared to the serious intentions, so we have established a set of criteria to evaluate each of the following playful factors: motivation, immersion, interactivity the gameplay and emotions.

So the objective of this element is to verify the efficiency of the playful mechanisms and factors offered by the

game from several angles, in other words if they help in immersion and involvement of the learner/player in the game.

2) *Playability Aspects*

Playability is characterized by different attributes and properties to describe the way in which the game can be played or the quantity that a game can be played. We are interested in two major factors or criteria that describe the playability aspects of the game: Replayability and Multi-player Value.

D. *Section IV : Technical Specifications*

Two major elements form this section in which we discuss the technical requirements and best practices for the game.

It is therefore to describe two major elements: first Technical Efficiency in which we check to see the efficiency of all the technical aspects of the game such as Graphics, Design, Sound, navigation ... secondly Requirements which aims to highlight all the requirements for the installation and start up the game.

1) *Technical Efficiency*

The first element of this section focuses on the technical aspects related to Game Design & Development in order to assess the effectiveness of each element that forms it. For this reason we have established a set of criteria to evaluate each specific side forming the game such as:

Graphics, Dialogues and sound, Browsing and Design. So through these criteria we first check the quality of the graphical aspects of the game if they are more realistic to ensure maximum immersion of the learner / player then we will evaluate the sound quality of the game to check if they are well chosen and placed in the game space. We also check if the game has a good design and provide a better navigation in the game since we know the good design makes always navigation easy (access to maps, training mission ...). In other words the choices in the game must be clear and grouped into coherent menu.

2) *Requirements*

To be used efficiently, all serious games need certain hardware components or other software sources to be well installed and used in the best conditions. These prerequisites are known and presented in different forms, some concerns just the installation process and others interested in improving the use of the game. They are often used as a guide and not an absolute rule. The requirements element covers a set of criteria such as: Install Requirements, System Requirements, Guide and Instructions in which we specify if there is any configuration or software installation required to implement the game as well as the system requirements (Hardware, Operating System, plugins, devices...) to run the game. We check also through these criteria if the game provides any kind of support (guide or Instructions) for the player/learner either at installation or during the use of game.

TABLE I.
THE ASSESSMENT GRID OF THE QUALITY OF SERIOUS GAMES

Assessment and analysis grid of the quality of serious games intended to be used in education field			
Serious Game Identification			
	<input type="checkbox"/>	Name of the game :	
	<input type="checkbox"/>	Game developer :	
	<input type="checkbox"/>	Game Type :	
	<input type="checkbox"/>	Game genre :	
	<input type="checkbox"/>	Age group Or ESRB rating :	
	<input type="checkbox"/>	Game experience required :	
	<input type="checkbox"/>	Game license :	
	<input type="checkbox"/>	Description :	
	<input type="checkbox"/>	Language :	
	<input type="checkbox"/>	Keywords :	
Pedagogical Specifications			
	Criteria	Rating (score/5)	Evaluation Factors
Content	Accuracy of the serious content		<ul style="list-style-type: none"> ▪ Consistency between the objectives and content of the game.[6] ▪ Identify any gaps or irrelevant content. ▪ Presence of structuring knowledge.
	Consistent with the target audience		<ul style="list-style-type: none"> ▪ Relevance of content based on audiences. ▪ Level of interest in the game for the target audience: the balance between learning time and fun time. ▪ Ability to edit issues (content and form)
	Degree of difficulty of activities		<ul style="list-style-type: none"> ▪ Degree of complexity of vocabulary (new words...). ▪ Presence of various issues.
	Feedback		<ul style="list-style-type: none"> ▪ Presence of a feedback mechanism in time related to each learning task to allow players to identify successful activities and those they have failed.

			<ul style="list-style-type: none"> ▪ Presence of motivational feedback messages offering encouragement and value learning made. ▪ Presence of review mechanisms content promoting feedback on the learning achieved in the game and access to support materials to review the learning that has not been made.[6]
Strategies	Specification of objectives		<ul style="list-style-type: none"> ▪ Are the objectives to be achieved stated and modeled as a list of skills to be acquired
	Existing Knowledge and representations		<ul style="list-style-type: none"> ▪ Are knowledge and existing learners' representations taken into account during the implementation of educational scenarios?.[6]
	Pedagogical Approach		<ul style="list-style-type: none"> ▪ Serious games are based on several important educational theories:[5] Constructivism. Socio-Constructivism.... So here we are checking if this serious game is based on one of these theories.
	Ability to adapt to each learner profile		<ul style="list-style-type: none"> ▪ Providing individualized learning based on different learner profiles, "training" adapts to users".
Assessment Method	Assessment procedure		<ul style="list-style-type: none"> ▪ Does the serious game provide an assessment procedure which take in consideration both the overall distance learning setting and knowledge acquired by learner / player?
	Progress indicator		<ul style="list-style-type: none"> ▪ Presence of progress Indicators which give information to player / learner about his current progress towards a closure in addition to the configuration of game elements involved
	Player Tracking		<ul style="list-style-type: none"> ▪ Does the game allow us to track the player/learner by the systematic collection and analysis of information during his current progress in the game? [8]
	Presence of tutor		<ul style="list-style-type: none"> ▪ Does the serious game allows the presence of one or more tutors during the game in order to creates additional pressure for the player/learner to try harder during the game and helps him to proceed to a more fruitful direction...
	Assessment Type		<ul style="list-style-type: none"> ▪ Does the serious game provide different type of Assessment: summative / formative individual / collective...?
Playful Specifications			
	Criteria	Rating (score/5)	Evaluation Factors
Attractiveness (Fun)	Motivation		<ul style="list-style-type: none"> ▪ Presence of a set of resources to ensure the players perseverance in the performed actions to overcome the game challenges. This means, different factors to make sure positive behavior in the interpretation of the game process, focusing the player on the proposed challenges, showing the relevance of the objectives to reach and reward for challenges, promoting the player confidence to face them and the pleasure to achieve them.[9]
	Immersion		<ul style="list-style-type: none"> ▪ Does the game really help immersion and involvement of the learner/player in the game?
	Interactivity		<ul style="list-style-type: none"> ▪ Presence of interactivity which helps to maintain active learner manually and intellectually.
	Gameplay		<ul style="list-style-type: none"> ▪ Presence of one of the two gameplay types: "Game-based", designed with stated goals to reach like Avoid, Match, Destroy. The second type is "Play-based", designed with no stated goals to reach like Create, Manage, Move, Select, Shoot, Write, Random.[7]
	Emotion		<ul style="list-style-type: none"> ▪ Presence of emotions in serious games which help to obtain a best player experience and leads players to different emotional states: happiness, fear, intrigue, curiosity, sadness... using the game challenges, story, aesthetic ...[9]
Playability Aspects	Replayability		<ul style="list-style-type: none"> ▪ Does the game allow the replayability which used to describe the entertainment value of playing a game more than once?
	Multiplayer Value		<ul style="list-style-type: none"> ▪ Does the game authorize more than one person to play in the same game environment at the same time?
Technical Specifications			
	Criteria	Rating (score/5)	Evaluation Factors
Technical Efficiency	Graphics		<ul style="list-style-type: none"> ▪ Through this Criteria we check the graphical aspect of the game if the graphics are more realistic and graphic quality of the characters, avatars, backgrounds ...
	Dialogues and Sound		<ul style="list-style-type: none"> ▪ It is important to check if the game offers better sound quality because the players appreciated even if it does not become their primary endpoint, so through this Criteria we check to see if the sound resources are well placed and highlighted in the game space.
	Browsing and Design		<ul style="list-style-type: none"> ▪ The game design should facilitate navigation in the game space. While manipulating the resource, the player should be able to load maps find a plan, mission or a detailed description of stage. The suggested choices should be clear and the groupings within the menus should be consistent.[10]
Requirements	Install Requirements		<ul style="list-style-type: none"> ▪ This Criteria describes whether there is software installation will be required or setup configuration to run the game.
	System Requirements		<ul style="list-style-type: none"> ▪ Does the game require a specific system configuration (Hardware, Operating System, plugins, devices...)
	Guide and Instructions		<ul style="list-style-type: none"> ▪ We wonder if the game provides support elements for the player/learner either at installation or during the use of game (Procedure, tutorial, Installation Instructions, explanation of rules...)

IV. EXAMPLE OF USE: SERIOUS GAME USED TO HELP STUDENT LEARN COMPUTER PROGRAMING

Before we apply the proposed grid we first try to explain how to use it:

A. how to use the grid

In order to check if every serious game meets the various criteria cited in the grid we have set up a scoring system that gives a score ranging from 0 to 5 on Likert Scale¹ for each criterion. The result of this notation gives us the opportunity to make two types of evaluation in the same time.

1) An Overall assessment

In which we make a global assessment of serious games intended to be integrated in the educational field. This assessment type that uses the rating method following Likert Scale¹ allows therefore having a general idea about the quality level of games assessed.

TABLE II.
THE APPROPRIATE QUALITY LEVEL DESCRIPTION FOR EACH SCORE RATING OBTAINED

Overall Score	Ability to be Integrated Into Education (%)	Quality Level
0 ⇒ 52	From 0% to 40%	Substandard: The game is far from being used in an educational context since it does not meet the criteria expressed.
53 ⇒ 78	From 41 % to 60 %	Average: The game can be used in education but it does not allow a sufficiently significant educational use.
79 ⇒ 104	From 61 % to 80 %	Superior: in this case the game is sufficient and adequate for educational use, although it does not allow a pedagogical exploitation at full potential.
105 ⇒ 130	From 81 % to 100 %	Very Superior: This level of score allows us to achieve an excellent level of quality for educational use since it meets the required quality criteria.

2) Specific Assessment

The first type of assessment gives an overall view of the quality of the game which prevents us therefore to assess each independent section in order to find the sides of deficiencies, unlike this second type of assessment that focuses on the evaluation and analysis of each section.

Each section is characterized by a set of criteria which each criterion will be reviewed by a score ranging from 0 to 5. So in order to evaluate each section simply divide the sum of the scores obtained on the total score of the section:

- Pedagogical Specifications ⇒ Sum of scores / 65
- Playful Specifications ⇒ Sum of scores / 35
- Technical Specifications ⇒ Sum of scores / 30

¹ Lickert scale is a rating scale where the interrogated person expresses his/her degree of agreement or disagreement regarding a statement. The scale contains five or seven answer choices which enable to formulate the degree of agreement.

Thus, this method allows us to verify through the score assigned if the game reaches the required level of quality in the Educational, Playful and Technical section.

B. Application of the Grid

In what follows we try to apply the proposed grid on an example of serious games used to help student learn computer programing:

1) Robocode

Robocode is a programming game, where the goal is to develop a robot battle tank to battle against other tanks in Java or .NET. The robot battles are running in real-time and on-screen. [11]

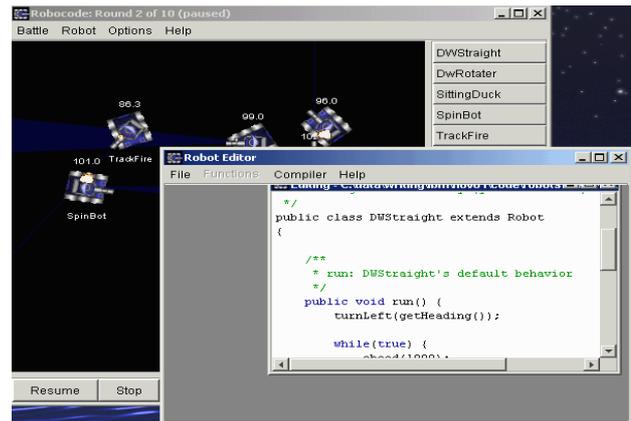


Figure 2. The Robocode IDE

TABLE III.
SERIOUS GAME IDENTIFICATION EXAMPLE (ROBOCODE)

Serious Game Identification	
Name of the game :	Robocode
Game developer :	Flemming N. Larsen (IBM)
Game Type :	Programming game
Game genre :	Action (Combat Sim)
Age group Or ESRB rating :	N.A
Game experience required :	Beginner
Game license :	Eclipse Public License (EPL)
Description :	Programming game, where the goal is to develop a robot battle tank to battle against other tanks in Java or .NET.
Language :	English
Keywords :	Robot, Tank, programming, Battle, battlefield

We started with an Identification of the Robocode game before we proceed to an overall assessment in order to get an idea about its quality level and the ability to be integrated into Education:

After applying the proposed grid on the game by assigning a score for each criterion we got an overall score of 67 which mean that Robocode has reached 51.5% of criteria, a percentage reflecting the average quality level of the game.

Regarding the evaluation and analysis of each section we obtained the following results:

- Pedagogical Spec. \Rightarrow Sum of scores (27) / 65 \Rightarrow **41.5 %**
- Playful Spec. \Rightarrow Sum of scores (21) / 35 \Rightarrow **60 %**
- Technical Spec. \Rightarrow Sum of scores (19) / 30 \Rightarrow **63 %**

2) *Synthesis*

So, as a synthesis of the two types of assessment carried:

- 51.5% allow to Robocode to be used in education but it does not allow a sufficiently significant educational use.
- Robocode has reaches 41.5% of pedagogical criteria which reflect a pedagogical need specifically in assessment method element.
- The score obtained for the playful specification that reaches 60% shows that the graphic areas and scenarios of the game are rather exciting but not attractive enough to keep the player's attention.
- For the technical side we scored a little above average (63%) which gives us an idea about the efficiency of all the technical aspects of the game and shows that the technical specifications of Robocode are sufficient for educational use.

We have applied our grid on other games used to help student learn computer programming such as: Prog&Play, Eclone, TA Spring, CoLoBot.

The result of the assessment of each of these games shows that each one has its own weak and strong points. But they all have unmet needs in the educational side especially the side that is interested in the assessment method adopted in the game also none of those games does allow tracking the learner/player during his progress.

V. CONCLUSION

Due to the lack of empirical studies on serious game and their integration with educational field and all the issues that come with such as motivate players, tracking the players, the balance between learning and playful aspects...it is imperative that more studies have to be conducted with different types of serious games, on different topics and targeted users in different contexts in order to ensure efficient and significant educational use.

In this paper we suggest an assessment and analysis grid which aims to assess the quality level of the serious games intended to be used in education. This grid that allows us to make an overall or specific assessment was applied on a set of serious games used to help student learn computer programming. As a result we were able to find the side of weakness or strength of each game assessed as well as its ability to be integrated into education.

As mentioned above all the assessed games have unmet needs in the educational side especially the side that is interested in the assessment method. So further effort is required in order to overcome the educational limitations of serious games.

REFERENCES

- [1] Yassine El Borji and Mohamed Khaldi. Article: An IEEE LOM Application Profile to Describe Serious «SG-LOM». International Journal of Computer Applications 86(13):1-8, January 2014.
- [2] Petros, P., & Georgios, F. (2011, June). Educational Applications of Serious Games: The Case of the Game Food Force in Primary Education Students. In Proceedings of the 7th European Conference on Management Leadership and Governance: Ecgl 2011 (p. 476). Academic Conferences Limited.
- [3] J. Alvarez. Du jeu vidéo au serious game, approches culturelle, pragmatique et formelle. PhD thesis, Université de Toulouse, 2007.
- [4] Abderrahim El Mhouthi, Azeddine Nasseh and Mohamed Erradi. Article: Development of a Tool for Quality Assessment of Digital Learning Resources. International Journal of Computer Applications 64(14):27-31, February 2013.
- [5] Kafai, Y. B.: Playing and making games for learning: Instructionist and constructionist perspectives for game studies. Games and Culture 1(1): 34-40, 2006. <http://dx.doi.org/10.1177/1555412005281767>
- [6] RYAN, William et CHARSKY, Dennis. Integrating serious content into serious games. In : FDG. 2013. p. 330-337.
- [7] D. Djaout, J. Alvarez, J.P. Jessel, "Classifying Serious Games : the G / P / S model," Medicine, 2008.
- [8] HALLINAN, Maureen T. Tracking: From theory to practice. Sociology of Education, 1994, p. 79-84.
- [9] LINDLEY, Craig A., NACKE, Lennart, et SENNERSTEN, Charlotte C. Dissecting play-investigating the cognitive and emotional motivations and affects of computer game-play. CGAMES08, 2008.
- [10] KOSTER, Raph. Theory of fun for game design. " O'Reilly Media, Inc.", 2013.
- [11] Li, S. (2002). Rock em, sock em Robocode. IBM developerWorks: Java technology-<http://www-106.ibm.com/developerworks/library/j-robocode>.

AUTHORS

Yassine El Borji is with Computer lab, operational research and applied statistics, Abdelmalek Essaadi University, Faculty of Sciences, Tetouan, Morocco (e-mail: elborji.yassine@gmail.com).

Mohamed Khaldi is with Computer lab, operational research and applied statistics, Abdelmalek Essaadi University, Faculty of Sciences, Tetouan, Morocco (e-mail: medkhaldi@yahoo.fr).

Submitted 02 September 2014. Published as resubmitted by the authors 15 October 2014.