Characterization and Evaluation of Serious Games: A perspective of their use in higher education

Imed Boughzala\textsuperscript{1}, Ikram Bououd\textsuperscript{1}, Hélène Michel\textsuperscript{2}

\textsuperscript{1}Telecom Ecole de Management
Institut Mines-Telecom, France
\{imed.boughzala,ikram.bououd\}@telecom-em.eu

\textsuperscript{2}Grenoble Ecole de Management, France
helene.michel@grenoble-em.com

Abstract

Serious games are video games or computer games designed for training or educational purposes. Thanks to the interesting and useful opportunities that they provide such as interactivity, immersion, simulation... they become well accepted and well spread among professional and academic fields. However, the selection of the most suitable serious games to a specific educational goal seems to be not well studied in the literature. This paper tries to meet this research gap by developing a new serious game evaluation grid basing on the design science approach. The grid is developed and tested with students and professionals.

1. Introduction

It has been a long time since digital games appeared in the ‘40s. They were used as research objects to study and illustrate human-computer interaction. Since 1946, with the MIT’ Whirlwind project, people witnessed the introduction of the first simulation that allows pilots of military aircraft to train in a controlled environment. Then, learning takes place by “trial and error” where flow and immersion are strengthened. The game Pong was the first public success in 1972. Later, with the democratization of video games, the use of simulators became popular, such as “Flight Simulator” which appeared in 1982. However, only recently the high potential of video games in the field of communication and knowledge transfer has been revealed. According to the ELM (Elaboration Likelihood Model), the model of persuasion proposed by [7], "a game can be seen as a playful approach to improve the motivation and perceived ability of individuals to process information cognitively" (cited in [10]). Few years later, this untapped aspect has created new class of games called Serious Games (SG). These latter are used as means to meet pedagogical goals such as training, simulation, education, promotion, communication etc. They are enrolled in the Environments for Human Learning (EHL) by combining machine-mediated learning, simulation, use of emotions and professionalism. Thanks to their approach combining seriousness and fun, the motivation and perceived ability of individuals to process complex information or to repeat behaviors have been improved [8]. Indeed, companies have moved to games for executive education in order to promote behavioral and inter-relational dimensions, not only for technical skills but also for general skills. Several institutional studies such as Scientists [15] and Tomorrow [17] confirm the idea that these games endow players with valuable skills for their curriculum and so transferable in the business world.

The increasing use of Information and Communication Technologies (ICT), and their continually rising complexity, offers a new horizon for training opportunities and tools [3, 15]. However, these technologies require that teaching methods would be adapted to the new generation of students (i.e. Generation Y) who increasingly use various and sophisticated technologies. Considering the transformation of pedagogy related to technology development and the momentum of active learning, the use of SG in higher education is enrolled in the desire of anticipation to be in line with the new learning methods and techniques already mobilized in colleges and high schools. In addition, using these advanced tools will prepare students to be familiar with them for the professional context in the near future. A SG is indeed a video game using the same design approaches and expertise as a classic ones but its playful approach goes beyond the single dimension of entertainment. It’s considered as a real tool for awareness, training or promotion. It is a "useful" kind of video games provided for professionals.
In recent years, the field of SG has exponentially grown and has been the object of numerous research studies. However according to our knowledge, very few researches have been focused on the evaluation of SG. The evaluation of their effectiveness, according to common criteria, would allow comparing them and establishing their legitimacy regarding other training means. The lack of reliable, reproducible and adaptable methods for evaluating SG, constitutes a research gap related to a real business need expressed by the professionals. Indeed, the increasing development of SG in several fields pose the problem of their characterization and so evaluation in order to facilitate their selection and adoption according to one pedagogical goal in higher education. Some researchers have suggested methods for characterizing and evaluating but they are offering limited view, specific to one domain or descriptive and non-reproducible.

This paper tries to meet this research gap and answers the following questions:

- Is it possible to develop a holistic approach as a reproducible grid to characterize and/or evaluate SG?
- How to determine which SG is more suitable for a given educational goal?

The objective of this research is to meet a business need following the Design Science approach to build grid for classifying and evaluating serious games. This grid, called G-CE-SG (Grid for characterizing and Evaluating Serious Games), was built and tested students and professionals.

The remainder of the paper is organized as follows: First, we present the literature review related to SG and their characterization/evaluation. Second, we focus on the method adopted in this research to build the G-CE-SG. Third, we present to application of this grid in higher education. Finally, we discuss the implications, limitations, and future research perspectives.

2. Background

2.1 Serious games

The term "Serious Games" seems to be an oxymoron. Indeed, according to the dictionary, a serious thing is a thing requiring thought, concentration, or application, as opposed to what is entertaining or distracting. Furthermore, the word game is defined as physical or mental activity, not imposed, not targeting any utilitarian purpose, and in which you devote yourself to have fun and gratification. So a game is an activity providing entertainment or amusement. Combining these two antagonist words “Serious” and “Games” one obtains a strong concept (serious games) related to a new class of games.

Alvarez (2007) [1] defines them as "computer applications having as original intention to combine both serious aspects […], with fun aspects from video games. Such an association is achieved by providing a learning scenario corresponding, from a programming point of view, to implement a decor (sound and graphics), story and suitable rules; therefore it moves away from restricting the game to entertainment". The main objective is to operate the entertaining aspect of video games to facilitate the learning of serious concepts which are traditionally taught through conventional teaching or training methods. The range of usage areas of SG is very wide such as scientific exploration, military, medicine or education. More precisely the educational aspect of SG is one of their greatest assets since they are promoting and opening new horizons for active learning.

Actually, according to [13] and [5], the pleasure and richness of experiences in the game will increase the learner interest to the topic. Thus, SG sublimate learning by making a learner interested in a subject that may not interest him/her from the beginning. Several studies have been conducted on SG in the education field. Thus, according to Rooney [14] one of SG’s challenges is to find the right synergy between pedagogy and learner engagement. Indeed, strongly favoring pedagogy at the expense of engagement risks to make learner loosing the interest in the game. Conversely, favoring engagement at the expense of pedagogy risks to make SG losing their original utility. Further research on the use of SG by children was also conducted. For example, Peteron [12] shows that SG help children to acquire skills and abilities such as: strategy, problem solving, logic, psychomotor coordination, concentration, motivation, organization, memory, creativity, exploration, communication, group work and decision making. In the same paper, the author emphasizes that research studies should be conducted before integrating the principles of theoretical education in games. Indeed, the majority of researches do not focus on the theoretical aspects necessary before the creation of SG. In order to address this issue and analyze in depth these games, one can evaluate them in order to characterize them and find the criteria on which their design is based. Consequently, one can use these latter to better customize games regarding their application field mainly those in higher education. However, SG’
evaluation is a research stream that remains largely unexplored mainly those targeting higher education.

2.2 Serious games’ characterization and evaluation

Two main grids for evaluating SG have been developed. The first one is G/P/S proposed by [4]. This grid allows evaluating (approximately) the gameplay of SG with gain comparing to SG without gain. In addition, it allows characterizing the serious objective of the SG (passing a serious message, training, advertisement etc) and to define the sector targeted by the game (military, health, public etc). Thus, this grid provides a classification scheme that could not be considered as an assessment tool. In addition, only three aspects have been taken into account namely gameplay, objective and targeted sector. Indeed, this grid didn’t provide any qualitative information, thus we cannot judge if a SG is effective, clear or easy to use. Finally, the number of evaluated criteria related to these aspects is restrictive (only six criteria).

Second, Peterson [12] proposed a grid allowing the evaluation of several aspects of SG for young children such as curiosity exploitation, mastery of the game, challenge, social, pedagogy, technology etc. This grid is the most complete grid found in the literature. Nevertheless, it refers to SG targeting children sector only. In addition, the value scale used by this grid is binary. Indeed, every aspect of a SG is divided into several criteria that the game could fulfill or not. Besides, this grid does not provide qualitative aspects which are necessary to achieve a complete evaluation of SG.

This grid is dedicated to educational SG only; it does not focus on other fields. This grid does not determine give any description of technological aspects (e.g. accessories, architecture, programming…). No information is provided on the game execution. Besides, the grid does not focus on users’ appreciation which can be useful for new users of the game. Finally, this grid has a high number of evaluation criteria (44).

So, there is a research gap to be studied, namely the development of a grid to evaluate a SG for higher education. This grid must be reusable and as precise as possible to cover all useful aspects for a complete evaluation. This grid should take into account different criteria omitted by the previous grids and especially to allow qualitative evaluation using a graduated measurement scale.

<table>
<thead>
<tr>
<th>Name</th>
<th>G/P/S Grid</th>
<th>Peterson Grid</th>
<th>G-CE-SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid objective</td>
<td>Classification</td>
<td>Evaluation</td>
<td>Characterization and Evaluation</td>
</tr>
<tr>
<td>Value scale</td>
<td>Graduated and binary</td>
<td>Binary</td>
<td>Graduated</td>
</tr>
<tr>
<td>Evaluated aspects</td>
<td>Objective</td>
<td>Objective</td>
<td>Objective and subjective (qualitative)</td>
</tr>
<tr>
<td>Audience targeted by the SG</td>
<td>Public Professionals Students</td>
<td>Children Children with special needs</td>
<td>Public Professionals (public and private corporation) Professional in higher education Students (initial and executive education)</td>
</tr>
</tbody>
</table>

Table 1: Comparison between evaluation grids

3. Method

The present research is based on a design science, a constructivist approach. If the behavioral science seeks to explain and predict phenomena that are related to organization’s business needs through the development and justification of theories (Applicable knowledge), the Design Science tries to meet the identified business needs through the building and evaluation of artifacts [6].

The design science approach is very suitable for the development and application of the G-CE-SG by demonstrating its practical feasibility and utility. Regarding the context of SG, the G-CE-SG is, actually, not only an empirical elaboration of a SG’ evaluation instrument, but also a starting point for further research in this area, and therefore an object for improvements and adjustments.

March and Smith [9] identify two processes and four design artifacts produced by design science research. Processes are the building and evaluation. First, we
will develop the grid following a design science approach and then we will evaluate them. This evaluation will bring many feedbacks and inform on the fit of the designed grid (set of artifacts) to the requirements of the business need. The evaluation of the G-CE-SG presented in the next section is an evaluation conducted with students from both a Business school (will be called after TEM) and Engineering school (will be called after TSP) hosted in the same campus in France. In addition, its artifacts are assessed later according to the seven guidelines of Hevner et al.’s [6, p.86] design evaluation framework.

The four artifacts related to the G-CE-SG are constructs, the model, the method and the instantiation. In this case, these artifacts would be represented as follows:

- **Constructs:** A construct is defined as the semiotic and symbolic field where the research is enrolled [6]. Thus, the G-CE-SG structure and the set of criteria used to describe the characteristics of SG.

- **Model:** The set of questions and measurements scales related to the criteria of the grid. The model is presented as questionnaire.

- **Method:** Refers to how to use the grid in terms of process, steps and supports a set of recommendations.

- **Instantiation:** the G-CE-SG tool which is a customized Microsoft Excel application that represents the implementation of the above artifacts, and enables the execution of a concrete evaluation.

4. The G-CE-SG

4.1 Description

The G-CE-SG aims to characterize and evaluate holistically a given serious game. It supports the development of recommendations in form of an action plan to streamline its usage for a specific educational objective regarding particular conditions. Its applicability is not limited to a particular field even if the initial business need was expressed for higher education. The grid can be used for different pedagogical settings and fields.

The last version of the G-CE-SG is divided in three sections: A first section which gives a brief identification of the SG. This section is called the identity card. The second section contains the set of criteria allowing the characterization of the SG. This section is called the descriptive card. Finally, a third section, with more subjective details through a set of qualitative criteria of the SG. This section is called the evaluation card.

The following table gives a synthetic view of the G-CE-SG:

<table>
<thead>
<tr>
<th>Identity card (entered by the user)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifier: …</td>
</tr>
<tr>
<td>Name: …</td>
</tr>
<tr>
<td>Web link: …</td>
</tr>
<tr>
<td>Cost: …</td>
</tr>
<tr>
<td>Editor: …</td>
</tr>
<tr>
<td>Edition date: …</td>
</tr>
<tr>
<td>Release: …</td>
</tr>
<tr>
<td>…</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Descriptive card (chosen by the user among multi value lists)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Languages: (V1…Vn)</td>
</tr>
<tr>
<td>Age:</td>
</tr>
<tr>
<td>Expertise Field:</td>
</tr>
<tr>
<td>Pedagogical purpose:</td>
</tr>
<tr>
<td>Requested level of expertise:</td>
</tr>
<tr>
<td>Game mode:</td>
</tr>
<tr>
<td>Personalization:</td>
</tr>
<tr>
<td>Number of simultaneous players:</td>
</tr>
<tr>
<td>Number of scenarios:</td>
</tr>
<tr>
<td>Accessories:</td>
</tr>
<tr>
<td>Support:</td>
</tr>
<tr>
<td>Evaluation mode:</td>
</tr>
<tr>
<td>Graphical display:</td>
</tr>
<tr>
<td>…</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evaluation card (chosen by the user among multi value lists)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easiness of set up and configuration: (V1…Vn)</td>
</tr>
<tr>
<td>Interface ergonomics:</td>
</tr>
<tr>
<td>Graphics richness:</td>
</tr>
<tr>
<td>Complexity:</td>
</tr>
<tr>
<td>Animation:</td>
</tr>
<tr>
<td>Cognitive load:</td>
</tr>
<tr>
<td>…</td>
</tr>
</tbody>
</table>

**Overall rating: (1…10)**

Table 2: G-CE-SG: Grid for Characterizing and Evaluating Serious Games.

4.2 The development process of the G-CE-SG

The process that allowed developing the G-CE-SG includes several steps and involves actors:
4.2.1 Working group with experts

The idea of developing the G-CE-SG comes to meet a business need expressed by the CCMP (Centrale de Cas et Médias Pédagogiques i.e. Cases and Pedagogical media collection publishing) of the Chamber of Commerce and Industry of Paris. The CCMP wanted to position itself as a reference in the domain by creating a quality label for SG editors and a guarantee for acquirers (the academia field actors) guiding them to make the best choice.

A working group was formed in December 2010 to meet periodically in order to address this issue. This group includes eight persons: four persons from the academia, three professionals and the Managing Director of the CCMP. The principle of the grid (the content) and the protocol of use (process) have been discussed from the first meetings. Then, participants come up with the idea to draw on quality maturity models, such as the Capability Maturity Model (CMM) [11], to build the grid with a list of criteria. The process for evaluating a serious game was defined including three steps: (1) the test of the serious game by an expert, (2) the application of the grid by the same expert for identifying, characterizing and evaluating the serious game, and (3) the experimentation of the serious game by a teacher in real condition within a classroom (tests with a group of students).

4.2.2. Brainstorming with students

In order to start thinking about the grid and its composition, a free brainstorming session has been conducted in March 2011 with students enrolled in the course of "Facilitation techniques for brainstorming and creativity" in TEM (Business school). The purpose of the brainstorming was: How to characterize and evaluate serious games through an instrument?

**Participants**

The following table summarizes the characteristics of students participating in the brainstorming session:

<table>
<thead>
<tr>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total of participants: 25</td>
</tr>
<tr>
<td>Grade Level of Participants: Master degree - 2nd year of TEM</td>
</tr>
<tr>
<td>Youngest Participant in Age: 20</td>
</tr>
<tr>
<td>Oldest Participant in Age: 23</td>
</tr>
<tr>
<td>Average age of participants: 21</td>
</tr>
<tr>
<td>Male Participants: 13</td>
</tr>
<tr>
<td>Female Participants: 11</td>
</tr>
</tbody>
</table>

**Table 3: Participants' demographic data**

**Process**

To conduct the brainstorming session a Delphi approach was used, enriched with thinkLets-based facilitation process contribution [2]. ThinkLets are codified best facilitation practices that create predictable, repeatable patterns of collaboration among people working toward a goal. They are used to streamline collaboration during brainstorming sessions, rapid decision-making, strategic objectives evaluation, team building, and creativity [17]. We had the opportunity to use a Group Support System (GSS) and a well-structured facilitation process to conduct the session. The brainstorming process consisted of several activities where participants were asked to contribute during a 120 minutes period. First, participants were asked to anonymously generate ideas around the theme “Characterizing and evaluating serious games through a reading grid”. Second, participants were requested to reduce, clarify and organize collectively the generated ideas into unique statements around the subject. The goal was to converge on similar ideas, remove non-related ones, and reword those that were insufficiently clear.

Third, the facilitator presented and explained to the group the selected statements. Then participants were asked to rate individually and anonymously the relevance of each statement on a 5-point Likert-type scale, with =5' representing a very relevant statement and =1' a least relevant statement. Finally, the voting scores were presented (proposal by proposal) to all participants in a ranking format to stimulate a discussion of the results, to allow the reformulation of proposals when necessary, to clarify ratings' standard deviations and to build a collective consensus.

The final results of this brainstorming are seven ideas that received the best consensus with highest means and smallest standard deviations. According to the participants, important objectives to build such grid are as follows:

1. The grid must be as rich as possible to provide maximum analytical elements.
2. The grid should be divided into themes including several criteria.
3. The grid must be intuitive and easy to use.
4. The themes of the grid should provide a rich support to characterize the game.
5. The criteria must provide a deep understanding of the game features.
6. The measurement scales of different criteria should be graduated.
7. The process for using the grid should be targeted in terms of educational objective.

4.2.3 First version of the grid and initial tests (TSP students’ project)

From the results of the brainstorming session, two students in the second year of TSP (Engineering school) have been assigned to work on the development of the first version of the grid (April-July 2011), and supervised by one researcher (co-author of this paper).

The development of the grid has been preceded by a literature review related to SG and related theories (the entertainment theory, cognitive theories, flow theory, motivation theory, task-technology fit theory...) and evaluation grids. Then, they came up with the first version of the grid.

This grid was then tested with five games by these two students (Being the Big Boss, Stop Disasters Game, America’s army, and Take Back Illinois). Then, it was presented to the experts’ working group in order to be validated after considering the initial modifications (unclear or redundant criteria, criteria order, subjective criteria ...). In October 2011, the grid was presented to the TSP students during the course “Collaboration technologies” for getting volunteers for its application. Six students have accepted to test each one a game at their convenience and to provide their feedback. The six games tested were BNP StarBank, Power of Research, Ma Cyber Auto-Entreprise, Reveal by l’Oréal, Stop Disasters Game, and CitéJob Recrut.

This step resulted in the development of the first validated but not achieved version of the G-CE-SG.

4.2.4. Second version of the grid and extended tests (TSP students’ project):

After that, four students in the second year of TSP (Engineering school) have been assigned to work on the development of the second version of the grid (from February to June 2012), and supervised by the same researcher. This grid has thus been amended several times. Several criteria have been eliminated and the grid has been deeply reorganized. This reorganization had two main goals: to make the grid more logically structured and create the three different sections for more readability (identity, characterization and evaluation). Along the reorganization, five criteria have been deleted. These criteria were redundant and added misunderstanding to the grid (i.e. themes vs. application fields; interactivity vs. interactive tools and game mode; test levels vs. evaluation mode and number of scenarios). Then, the second version of the G-CE-SG was presented to the experts’ working group in order to validate the last changes for starting extended tests with the students of TEM (Business school).

4.3. The using process of the grid

The using process of the grid (the G-CE-SG method) includes five main steps:

1. Purpose scoping: Defining the educational objectives before the serious game evaluation.
2. SG testing: thorough test of the SG.
3. Grid filling up: Using the G-CE-SG and answering/scoring the criteria.
4. Training experimentation: Determining of conditions of the training provided with the game and feedback from learners (game users).
5. Training disposal validation: Writing a user guide and recommendation for the SG.

5. The application of the G-CE-SG within a Business School

Students enrolled in the course of "Virtual Worlds and Serious Games" in TEM (Business School) were invited to participate to a large application of the second version of the G-CE-SG. The field application followed the G-CE-SG method steps (§ 4.3). It was performed between April and May 2012. As presented to students, the objectives of this application were to check:

- If the serious games are suitable to be used at a Business school (and in higher education in general) as a new alternative for learning.
- If the grid is well done to characterize and evaluate a serious game easily.
- If there are some improvements to the grid.

24 students have accepted to participate to this application. All of them are familiar with video games and have already tested at least one serious game before. These students were asked to test a given game for few days. A rich set of games were proposed to them by the researcher in which they could choose according to their preference. Then, the G-CE-SG questionnaire was sent, through an electronic form, to the students for fulfillment. The data collection and
analysis was done thanks to the G-CE-SG tool. Few open ended questions were added for getting participants’ feedback.

The G-CE-SG was applied for 24 serious games proposed on the market such as Darfur is dying, Global Challenge, Staying Alive, Action police, Energyville, Alcootel, Moonshield, Trading 212 demo, Peace corps, Cesim, etc.

The application of the grid has showed relevant and useful information for deep understanding of the tested games characteristics and qualities. In addition to characterize each SG, this helps to classify games in terms of expertise field, type, sector, purpose… and others proprieties related more to the judgment of each person: ease of use, perceived usefulness, interface ergonomics, complexity, etc. Some of these SG were tested by more than one student. The comparison of responses (especially the evaluation section) is in some cases very interesting.

Here after, some diagrams to present the results of the grid application are provided. These results show that the SG tested by students cover various fields, 14% for education and 22% for management (Figure 1). 48% of these games are edutainment games (Figure 2). 44% of the pedagogical purposes of these SG are for learning (Figure 3). 39% of them have a Strategy-builder as pedagogical model (Figure 4).

In addition, the results show that 43,5% of customization of these SG is at the player/avatar level. 60,9% contain no interaction tools between different players (gamers).

Students noticed also that 34,8% of SG contain only one scenario while 43,5% of them provide more than five scenarios. Besides, 52,2% score the player during the game execution while 60,9% give the evaluation at the end of each scenario.

Results show that 56,5% of these SG are in 2D technology and 95,7% of them are remotely downloadable or free. Actually, participants noticed that 34,8% of tested games are free with copyright and 91,3% are played on a PC / Mac / Netbook while none is requiring accessories.

In addition, students found that 69,6% of the tested SGs very easy to set up and configure while 52,2% of them have good interface ergonomics. Only 39,1% have a good graphics richness and only 43,5% have a clear purpose. In terms of complexity, 30,4% of them are considered as very simple and having basic animations. Finally, 65,2% of participants like the games they tested and give a good overall rating (Average score: 7,34).
In terms of feedback related to the use of the G-CE-SG to characterize and evaluate SG, some interesting assertions were provided by students, such as:

- The grid provides very rich information.
- The use of the grid requires going further in understanding the features of the game.
- Interface ergonomics have an important impact on the user attitude towards the game usage.
- The possibility of interaction with other players / elements of the game environment is important.
- Several terms (i.e. criteria names, measurement scales) remain unclear, such as cognitive load and remote use.
- Allowing the respondent to check more than one choice for some criteria (e.g. age categories, game type…).
- The scoring system should be reviewed for some criteria.
- Some criteria should include open choices.
- Some criteria seem very technical in nature, sometimes are not clear for everyone, and or where the information is not available.

6. Discussion and implications

In this paper, we followed the seven guidelines for design science as proposed by Hevner et al. [6]. These guidelines help to check to which extent the G-CE-SG gives an effective and efficient response to the initial business need. Indeed, this means that this grid, as an artifact, acts not only on a technical level, as it serves as a means to characterize/classify/evaluate SG basing on their characteristics and qualities, but also it explores the complex subject-object relationship when considering the SG in its micro and macro-social environment (real conditions of use).

In order to produce new artifacts [G-CE-SG structure, questionnaire, method, and tool] to be added as applicable knowledge to the knowledge base [see IS research framework in Hevner et al. [6, p.80], we developed a purposeful instrument and application (guideline 1: design as an artifact) showing step by step how to use the grid and solve a specific problem related to the overall assessment of a SG. This managerial instrument meets a business need clearly expressed by professionals and society (through the evolution of educational system and means), namely: the difficulty to choose the most appropriate SG that suits with a specific educational goal (Guideline 2: Problem Relevance).

The G-CE-SG not only allows to better detect the (internal) characteristics of SG, but also evaluate their impact and interaction level with the user. 24 SG were assessed according to the grid, 24 students participated in the application of the grid to assess its relevance and utility. Besides, an active contribution of a working group of experts (through a focus group) has been established to validate the G-CE-SG (Guideline 3: Design Evaluation).

Our literature review showed that the SG evaluation seems to be not well studied in the existing research. Mostly two grids were developed before, but they are either very simplistic to rigorously meet the business need (a characterization / classification more than evaluation), or very complicated to apply. In addition, the involved experts have clearly highlighted the business need to create an easy to use and rigorous assessment instrument (Guideline 4: Research Contributions).

Several work sessions with three different actors (TEM and TSP students and experts) have been conducted, thus a first version of the grid has been developed, then a second version has been proposed to improve the quality of the grid (Guideline 5: Research Rigor). In order to develop the grid, we used a combination of methods and researches including a literature review, a working group of experts, and Lab studies (Guideline 6: Design as a Search Process).

Finally, our results are and will be published in two stages (Guideline 7: Communication of Research): First, the G-CE-SG (with the process of use) and training experimentation will be presented through publications to other researchers who, we hope, will consolidate and extend the grid and application, and to experts who could apply the method and provide feedbacks and recommendations for its future enhancement. The CCMP will communicate on it in the field, which will accelerate its diffusion and legitimacy. Second, after a further and deep study of the grid and its application in various contexts with SG, researchers could reuse it as a background to improve it and / or develop new instruments.

7. Conclusion

In this paper, we report on the development and a field application of a new evaluation instrument for SG. It was developed in an inductive perspective to meet a real business need as expressed by the CCMP-CCI. Our contribution is both theoretical and practical as we propose a grid, an application process, a supporting tool, and empirical application. The results should be of interest to academic researchers and higher
education practitioners interested in the use of serious games in their educational programs. The research contributes to serious games literature, theory and practice through the development of G-CE-SG artifacts that provide evidence of proof of value and proof of use in the field.

Nevertheless, there are limitations related to this work in order to complete the design science evaluation framework. First, our empirical evidence is based on only a single and partial field application done by students. Future applications should be done by educators and trainers. The G-CE-SG method is not completed since only three of the five steps are performed. Further field studies have to be executed to expand the evaluation of the G-CE-SG artifacts and to further enhance the G-CE-SG current version. Particular care will have to be taken to ensure that G-CE-SG can determine conditions of the training provided with SG and bring recommendations for the teachers. Second, at this stage, the G-CE-SG cannot yet be used to investigate a correlation between the quality of the serious games and education/learning goal. However, it provides a first step into this direction. In the future, we will focus only on the 3D serious games.

8. References


852