Individual and collaborative Performance and Level of Certainty in MetaVals

[1]Mireia Usart, [1,2]Margarida Romero
[*]1ESADE Law and Business School, 60-62 Av. Pedralbes, Barcelona, Spain,
mireia.usart@esade.edu
2Université Laval, Québec, QC G1V 0A6, Canada, margarida.romero@esade.edu

Abstract

In the context of Higher Education (HE) in general, and management education in particular, the use of Serious Games (SG) is spreading, and solutions are increasingly developing. Nevertheless, the implementation of this learning methodology deserves further study concerning pedagogical and psychological aspects such game performance and players' metacognitive processes. This paper aims to study the relation among these two variables, based on the review of the results of MetaVals SG during the last 3 years. MetaVals is a collaborative, computer-based SG designed to facilitate collaboration and metacognitive awareness among HE students. It has been played by 250 students in 16 different experiences since its first version, in 2011. Overall results show a high performance for collaborative phases of the game, furthermore, students' elicitation of their Level of Certainty (LC), although not significantly, could be related to a better performance. These results can be a basis for further studies focused on the implementation of collaborative GBL in formal and informal adult learning contexts. However, some challenges are also identified and discussed on the present version of MetaVals game, and solutions are proposed in order to continue with the design of SGs for wider application and learners' needs in the diverse actual contexts.

1. Introduction

The use of Serious Games (SG) in Higher Education (HE) has been increasing during the last decades. This implementation of games in the curriculum for educational purposes has led to the emergence of Game Based Learning (GBL) methodologies; especially in computer-based and online contexts [1][2]. Due to this fact, the use of games for educational purposes has been an increasing focus of interest for instructional designers, teachers and researchers, and initiatives such as the European Project Games and Learning Alliance (GaLA NoE) are a testimony of it. Present trends in HE are focused on student-centred, active learning models including SGs. In particular, we can affirm that Game Based Learning (GBL) has long been used in management education in order to help practicing skills and competences. According to different authors [3], GBL could play a central role in students' training in general and in adult learners in particular.

A particular case of SGs application is collaborative Game Based Learning (GBL). Collaborative GBL can be considered a powerful educational technique aimed at enhancing collaborative learning [4]. This educational instrument allows a realistic collaborative "learning by doing" approach that avoids real-life risks [5]. According to Herz [6], multiplayer games can transform knowledge into social capital because peer acknowledgement appears to be a powerful incentive for students. The analysis of the students' learning performance in relation to metacognitive aspects has been a wide field of research in collaborative learning [7] in general, and in collaborative GBL in particular. Students collaborating in groups must be able to monitor and adapt their cognitive and metacognitive processes to the changes in their motivational state, and determine how much social support may be needed to perform a task [8]. Despite the advantages of collaborative GBL in terms of motivation [9] and development of cognitive skills, some potential shortcomings are recognised and must be faced, as we will further explain in the next sections.

Starting from the study of the (meta)cognitive aspects of collaborative GBL, we will focus on the learners' Level of Certainty (LC) in relation to the learning process [10]. As we will further explain in the next sections, previous results have shown that LC elicitation should force students to be more aware of their metacognitive process and, consequently, show a higher learning performance [11][12]. Nevertheless, little has been said of this relation in the particular case of GBL activities in HE [13].

The general aim of this study is to review and summarize the results of MetaVals gameplay sessions during the last 3 years (2011-2013), since the SG was firstly designed. We study students' metacognitive monitoring; in particular, we focus on the Level of Certainty (LC) that is related to students' individual and collaborative performance in the game. We will analyze this relation, both for individual and collaborative stages of the game, in different European HE contexts. The MetaVals study has aimed to advance in the analysis of metacognitive supports in SG, focusing on the impact of supporting the students' metacognitive judgement of their Level of Certainty (LC), in the context of a collaborative GBL task. This 3-years research project has been organised around two research questions: 1) Collaborative learning will show better results in performance than individual learning. 2) LC elicitation will improve learners' performance in GBL, especially, in the case of collaborative performance.

These research questions are operationalized into three different hypotheses. The first one compares individual and collaborative performance:

H1: Based on previous results for Computer Supported Collaborative Learning (CSCL) in general, and on the particular case of GBL [13], we affirm that individual performance is, in general, lower than collaborative performance in a GBL environment.

The second hypothesis can be divided into two statements, relating performance with LC elicitation. In particular:

H2: We predict that both individual (H2a) and collaborative (H2b) game performance are positively correlated with students' LC elicitation. This hypothesis is based on previous studies observing the pedagogical benefits of a metacognitive support of the learners' interaction with the SG and showing a higher performance in the SG, both for individual and collaborative phases, when students elicit and share their LC [12].

In order to study these hypotheses, we will first summarize the main aspects of GBL in HE, to further focus on collaborative GBL. Second, we will explain the concept of Level of Certainty (LC); third, we will focus on the relation between students' LC elicitation and performance in collaborative GBL. Testing the hypothesis related to the individual and collaborative game performance in relation to LC requires the use of a collaborative SG supporting the expression and sharing of the LC. The need for this specific SG has leaded the authors to design and develop a SG called MetaVals to conduct this 3-year analysis. The SG MetaVals allows playing in in teams of two players (dyads) and introduce a metacognitive tool supporting the LC expression and sharing.

2. Collaborative GBL

GBL has long been used in HE for training different skills and competencies [15], through the use of a student-centred learning approach. Furthermore, as studied by Moreno-Ger and colleagues [16], the mix of fun and learning introduced by the GBL methodology could neutralize some of the negative learning outcomes and engaging expected from the results in classic learning activities. The use of GBL, nevertheless, has sometimes lacked of research measuring the learning effectiveness of these educational tools in general [1] and centred on the learner's reflection on the cognitive process in particular. Focusing on game performance, there are some challenges that must be taken into account; we should study if the implementation of tools that allow students to reflect on and share their cognitive level could help in their learning performance. Furthermore, collaborative learning activities in general have been approached as good contexts for learning, due to the fact that these contexts include a variety of educational practices in which interactions among peers constitute the most important factor in learning [17], but little has been said in the field of collaborative GBL.

2.1. Collaborative GBL in HE

Collaborative GBL can be defined as a particular case of collaborative learning. Following Wendel, Gutjahr, Göbel and Steinmetz (p. 287) [18], "the combination of game-based learning concepts and

collaborative learning may enable new, game-based application areas of CSCL, like collaborative multiplayer Serious Games". In particular, as can be seen in Figure 1, Computer-based collaborative GBL is an interdisciplinary field of research which should be analysed within the context of the GBL and the Computer Supported Collaborative Learning (CSCL) approach [13].
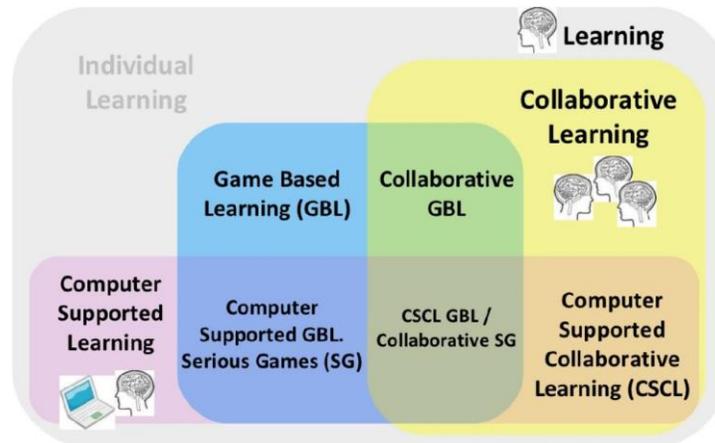

Figure 1. Collaborative GBL in the context of CSLC [13]

When compared to individual learning tasks, CSCL is supposed to enhance peer interaction and work in groups due to the fact that collaboration and technology can facilitate sharing and distributing of knowledge and expertise among community members [19]. Results of studies on collaborative versus individual activities in CSCL point to the fact that cognitive processes, needed for deep learning and information retention, may occur in dialogues [20]. Furthermore, CSCL effectiveness is influenced by external factors, such as the learners' prior experience, background, availability and expectations, the capacity of technology and technology-based delivery and other resources [21].

Focusing on GBL, and as stated by Prensky [22], for an educational game to be effective in terms of learning, its design must achieve a balance between fun and educational value. In this respect, collaborative GBL has the potential to enhance collaborative learning by designing an intragroup dynamic of cooperation between team-mates and a dynamic of competition to keep the gameplay engaging [23]. A collaborative GBL activity can also be played in different situations, from on-site classroom based situations to distance learning environments, with players geographically distributed and interacting in real time or asynchronously. Furthermore, both in onsite and online contexts, the role of the teacher or guide in the activity could be a key point for the success of the learning task. In particular, there are SGs that involve a virtual tutor or a teacher who guides the activity from the outside, especially focused at the previous and post discussion on the activity [24]. SGs can be performed to underline the alternative to traditional teaching methods and materials, even if the system under lens presumes some more strict rules of use with respect to the learning context. The use of SGs can be set as a self-evaluation activity from the students' perspective and as a formative assessment from the teacher's point of view, as we will see in our case study. Despite the pros of Collaborative GBL compared to individual gameplay, there are some challenges that must be kept in mind. It is a fact that some learners may have preferences for individual learning [25]. Furthermore, interactive environments such as games, simulations and adventures, sometimes lack effectiveness when no instructional measures or support are added to guide the learning process [4].

Students collaborating in small groups must be able to monitor, judge and further adapt their (meta)cognitive processes in order to perform the task, helping them to regulate both individual and collective actions in CLE [7]. Nevertheless, it is quite possible, following Munneke and colleagues [26], that arguing does not lead to more understanding of the issue, for example; when people stick to their own viewpoints, or peers do not advance with very strong arguments. Therefore, a necessary condition for successful collaborative GBL is that the game environment must support interaction by more than one player. Multiple players in the collaborative GBL situation may be a part of a minimal dyad or even a large player community playing a Massively Multiplayer Online Game (MMOG) [27],

but they must have enough communication tools that support their interaction and learning process. In our study, we will focus on dyads as an example of small group collaboration.

Based on the previous results of CSCL in general; and collaborative GBL in particular, in our first hypothesis we will study if collaboration is a necessary and sufficient condition for a better performance in GBL, when compared to individual gameplay, in MetaVals.

3. The Level of Certainty (LC) in the learning process

Azevedo [8] affirms that students collaborating in groups must be able to monitor and adapt their cognitive and metacognitive processes to the changes in their motivational state, and determine how much social support may be needed to perform a task. Furthermore, Schraw [11] defines different variables related to metacognitive monitoring; in particular, he explains how to measure the degree of accuracy of students' judgments on their performance before, during and after a learning task. In order to study the metacognitive process of students in collaborative GBL, we will focus on the Level of Certainty (LC), also known as Level of Confidence (LC) or Confidence [11][28]. It can be defined as the metacognitive judgements of the learner when she/he evaluates how certain she/he is about the accuracy of her/his answer. For our study, we will not only evaluate collaborative LC, but also focus on individual LC elicitation.

Prior studies have related LC with performance in different learning contexts. Previous research has related LC elicitation to a higher "solo" performance, based on the premise that individuals reflecting on their own cognitive process could be aware of their knowledge, better represent what they know, and thus reach deeper learning results [29]. Valdez [30] studied this relation in test-like situations; he stated that students' confidence estimates could influence their ability to accurately control and adjust their responses to test items. In particular, Valdez states that overconfident students (individuals showing low performance and high LC) may more easily be drawn to multiple-choice selections that are near approximations of the correct answer. On the other hand, underconfident students (that is, students with a high performance but low LC) could spend unnecessary time on a few test items and thereby limit their opportunity to execute test-taking strategies. We can believe that students are self-reflecting on their learning process while making explicit their LC. This variable could therefore be considered as a metacognitive judgement evaluating the accuracy of an answer that helps the learner to reflect and monitor on his own learning process and performance.

Nevertheless, (meta)cognitive judgements do not occur spontaneously in all the learners during the learning process. Learners show different levels of metacognition awareness and metacognitive process [31][32], which results on the fact that LC cannot be assumed during the learning process, neither the elicitation of the LC in collaborative learning situations when learners must solve a joint problem together. In order to ensure that students make an elicitation of LC, the scripting of the learning activity should be especially designed for this purpose. In this context, Swartz [33] affirms that LC elicitation in multiple-choice tasks could help leveraging the simplicity of these learning activities, and thus provide useful information to both the student and the evaluator. Results of his study show an equal acceptance among learners when compared to not LC eliciting activities, and a higher accuracy in the test performance.

In the context of this case study, the computer-based collaborative GBL aims to give students a useful tool for eliciting and sharing their LC. As we have previously stated, LC elicitation should force students to be more metacognitively aware, both for individual and collaborative situations. This LC elicitation has been previously related to better learning performance [10] and to a higher number of subject-focused comments in game conversations via chat tools during the gameplay [13]. We consider the interest of LC elicitation at two levels. Firstly, in an individual situation, the learner who reflects on his own LC is provided with a metacognitive judgement that could help him increase his metacognitive awareness on the LC and finally allow him to better regulate his learning process. Previous research in this context show that individuals do not always reflect on their learning process; Merenluoto and Lehtinen [10] present a study focused on learners' awareness of the quality of their previously acquired knowledge when asked for new contents. The authors highlight that, if a learner does not see or understand any reasons for change, he/she will tend to ignore it rather than revise his/her prior knowledge. However, individuals with a low LC are more open to review their knowledge, compared to high LC students, who were not willing to do a conceptual change, and therefore use less metacognitive regulation strategies [34]. Thus, we can affirm that giving learners a

specific context for individual reflection could help them in their learning process and performance through motivational strategies and reflecting on their cognition.

Secondly, focusing on collaborative contexts, a cognitive conflict could appear when the answers from the different team mates do not coincide. This situation could finally lead to a decrease on the students' LC [35]. Nevertheless, this conflict could be allowing a better preparation for the conceptual change, and therefore, for significant learning. Furthermore, an inexistent expression of LC could lead peers into a poor feeling of community, due to the lack of partners' cues, and therefore end up in superficial learning outcomes. Different tools in collaborative GBL contexts (as a specific case of CSCL) are necessary to foster and support argumentation among peers, in order to help them achieve a deeper understanding of a subject and a better construction of knowledge [29]. According to Munneke and colleagues [26], there is a need for enhancing metacognitive processes in CSCL, in order to lead to better learning outcomes. One of the tools that could foster these processes is Knowledge Group Awareness (KGA) explicitness [36], which includes the expression of the Level of Certainty (LC) to facilitate KGA.

In the next section we present the MetaVals SG, a classification game that implements a LC tool with the aim to help both individuals and dyads during the gameplay, to reflect on their cognitive process and have a positive impact on the final performance of learners.

4. The MetaVals SG case study

In this section we focus on MetaVals SG as an example of collaborative GBL in HE, using a LC tool for helping students make elicit and share their LC when classifying each item of the game. We will further analyse the data from the different contexts where the MetaVals SG has been implemented in order to study the individual and collaborative performance, and relate it to LC elicitation.

MetaVals is a computer-based collaborative SG that has been adapted from a previously existing class activity used to practice basic finance concepts [13]. Despite the pedagogical interest of the initial face-to-face activity, the classroom time limited the number of students actively participating in this learning action, and therefore it was difficult to incentivize discussion among peers in this context. A computer-based environment allows engaging an unlimited number of students in the activity and engaged them in dyads to discuss their knowledge. A GBL activity was designed based on a first, paper-based version that leaded to an ICT-based version that was tested in different learning environments [13]. The present release of MetaVals is web-based, and can be considered as a classification game, with a first individual phase where students have to classify 6 items into different categories, and two collaborative phases with 6 different items: the correction phase, for correcting peer's answers if needed, and the discussion phase, where dyads have to reach consensus on the final answers of the 12 total items. Collaboration and competition elements are implemented in these two latter phases, where students must collaborate with their peer in order to win the game against the rest of the class. In order to facilitate online interaction [37], MetaVals has a virtual dyad version that allows players to play and interact in an asynchronous context. Furthermore, in order to foster the engaging competition element [23]; researchers implemented a classification dashboard in the final screen, where students could access their scores and compare them to the rest of the players. The sharing of prior knowledge and experience is also faced in MetaVals with the use of a pre-test on the content, and a first screen where players can self-declare their prior knowledge and experience (see figure 2).
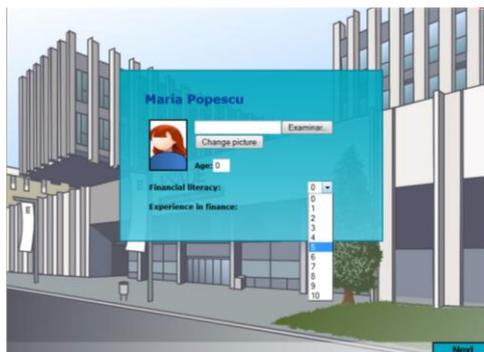
Figure 2. Screenshot of the MetaVals initial screen

The personalization of the game character aims to help players feeling identified with their avatar and also share significant information with their peers during the game. Finally, the MetaVals game implements a LC tool that evolves from the first, traffic lights tool [13] and aims to help students collaborating in dyads, as seen in the previous section, who must be able to monitor and further adapt their (meta)cognitive processes to possible changes in their motivational state. The present LC tool is a 10-grade scale accompanying each item:
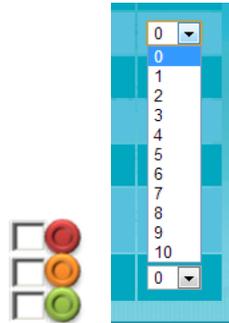

Figure 3. Evolution of the LC tool in the different MetaVals versions

In the following sections we will analyze individual and collaborative performance and LC results for the three year implementation of MetaVals in different contexts. In particular, this study focuses on the experiences in a Catalan School of business and Law (ESADE), in particular, there were eight master and executive master programs: Program for Management Development (PMD), Bancomer, MCDGE, ESADE internal training program, two Executive Master Marketing and Sales programs (EMMS and EMMV), Corporate training, and the Master in Marketing,; there was one context in the Universitat Oberta de Catalunya (UOC), one experience in the Universidad de Alicante (UAL) two groups in the University of West Scotland (UWS), and also two courses participated in Carol I university, in Romania (see table 1) This study was developed within the context of the Network of Excellence FP7 Games and Learning Alliance (GaLA), and in particular, within the context of the Special Interest Groups of Pedagogy and Psychology.

5. Method

5.1. Participants

This case study was carried out in 16 different contexts; in particular, MetaVals was implemented in three Spanish universities, an English university; and a Rumanian HE institution. Furthermore, it has also been played in two conferences: GaLA Alignment School 2011 and Online Educa 2012. A total of 250 participants with ages ranging from 21 to 49 years (M=30.05, SD=6.84), 127 men and 123 women played MetaVals in paper based (N=44) and Computer-based version (N=206). Nevertheless, four students in computer contexts only accessed the first individual phase and will not be taken into account for the analysis, finally, sample size was N=246. Students were Spanish, English and Rumanian. Players were mostly engaged in introductory finance courses (N=173), and therefore played the assets and liabilities version of MetaVals; in particular, there are 11 students from the two experiences in Carol I university that played this version with the objective of practicing English terminology. 47 students completed the learning theories version, and finally 19 players completed the Introduction to statistics version in UWS and Online Educa Berlin. LC elicitation was conducted in most of the contexts, nevertheless, in order to study the possible differences in performance due to CL, there were 6 contexts where half of the participants (as control groups) did not have access to the LC tool, in particular, 46 students did not had the LC tool available in their gameplays, and 200 students played with the LC tool (see table 1). The first results of the use of the MetaVals lead the authors to observe the advantages of introducing the LC tool. Because of this, the latest data collections using the

MetaVals has been conducted with the LC tool, leading to a higher number of participant having played the MetaVals game with the LC tool, and a lower number of participants in the control group.

5.2. Research design

To study our hypotheses, a quasi-experimental design was implemented in all 16 contexts. The use of a pre-test with 3 questions on content literacy, together with the collaborative GBL activity (MetaVals in its different versions) and a post-test on the experience of the game compose the case study scenario. MetaVals was played in its paper version, in face-to-face classroom with one computer for each participant, and finally, when dyads are virtual students can play at home, accessing the SG online in its present version. Finally, for testing H2, the first 6 contexts had a control group where students did not have access to the Level of Certainty tool, and a group with LC elicitation.

Figure 4. Screenshots of the MetaVals serious game, for the individual and collaborative phases

As we have outlined in the previous section, MetaVals consists of three different phases. In the first phase, players individually classify 6 items (phase 1) and rate their confidence for each item (see figure 5); this gives us the result of Individual Performance and Individual LC. Second, there is the first collaborative phase (phases 2), where students have to correct their partners' answers to 6 different items (Correction performance and LC). Finally, in phase 3, there are two different screens, the discussion self and the discussion dyad, where the two members of the dyad discuss and give a final answer for the 12 items. In this final phase we can measure Discussion performance and LC. It is important to note that, due to the fact that, when playing without a guide, most of the students do not perform the second discussion dyad screen; we will focus on the data from the first discussion screen in order to analyze the discussion phase LC. The complete game experience lasted about 35 minutes in average. As the participants finished the game, they were invited to fill the Post-test about the game experience, about the usefulness of the game, based on the Technology Acceptance Model (TAM).

Concerning the different contexts, the implementation of the items and content questions followed the same structure in order to maintain these contexts comparable. In particular, the difficulty of each item was placed in the same place (difficult for items 3 and 5, easy or medium for items 1, 2 and 4); and the pre-test questions were three multiple-choice questions also comparable among contexts.

5.3. Instruments and materials

Learning performance and Level of Certainty (LC) are measured using the MetaVals game. In particular, individual performance and LC are retrieved from the first phase of the game. Performance can have a maximum scoring of 6, if all the items are correctly classified, and LC a higher scoring of 10, when the student is totally sure of his/her answer. For collaborative results, the variables are retrieved in the second and third phases of the game, following the same process. The present version of MetaVals implements a MySQL database in order to monitor and record all the participants' actions, including LC use and both individual and collaborative scores, operationalized as learning

performance. The dyad which performs less errors on classifying the items, wins the game. Time spent in each phase is an untying factor.

6. Results

First we present a descriptive summary of the different contexts. In table 1, data for each context can be accessed. In particular: MetaVals version, use of the LC tool and its version, number of players, type of dyads and individual and collaborative average scorings for each group.

Table 1. Contexts data for MetaVals versions from 2011 to 2013

Context (N=16)	MetaVals version	Players	LC elicitation	LC tool	Dyads	Individual Performance		Collaborative Performance	
						M	*SD*	*M*	*SD*
ESADE 2011 PMD	Paper-based	28	16	Traffic light	Real	4.61	0.84	4.93	0.81
ESADE 2011 Bancomer	Paper-based	16	8	Traffic light	Real	5.16	0.92	5.65	0.58
ESADE 2011 MCDGE	Computer 1.1	18	10	10-grade scale	Real	4.55	1.01	3.81	1.75
ESADE 2011 Training	Computer 1.1	8	6	10-grade scale	Real	5.17	0.90	5.25	1.11
GaLA 2011	Computer 1.1	5	5	10-grade scale	Virtual	4.00	1.22	6.00	0.00
ESADE 2011 EMMS	Computer 1.1	25	13	10-grade scale	Both	4.25	0.86	4.10	1.56
ESADE 2012 EMMV	Computer 1.1	15	11	10-grade scale	Virtual	5.01	0.93	5.33	1.35
ESADE 2012 Corporate	Computer 1.5	17	17	10-grade scale	Virtual	5.47	0.80	5.47	0.94
UAL 2012	Computer 1.5	6	6	10-grade scale	Virtual	3.83	0.75	3.25	1.41
Carol I 2012	Computer 1.1	5	5	10-grade scale	Real	4.40	0.55	4.60	0.89
UWS 2012&13	Computer 1.5	11	11	10-grade scale	Virtual	2.92	1.62	4.00	1.53
Online Educa 2012	Computer 1.5	6	6	10-grade scale	Virtual	2.25	1.60	3.50	1.76
ESADE 2013 Marketing	Computer 1.5	16	16	10-grade scale	Real	4.53	0.57	4.94	0.57
UOC 2012	Computer 1.5	47	47	10-grade scale	Virtual	2.64	0.85	2.68	1.16
MOOC 2013	Computer 1.5	17	17	10-grade scale	Virtual	4.24	1.09	4.56	1.63
Carol I 2013	Computer 1.5	6	6	10-grade scale	Real	4.33	1.03	4.83	0.98

In order to study the first hypothesis (H1), that is, the relation between learning performance and context (individual or collaborative), an Analysis of Variance, or One-Way ANOVA was used. As

previously explained, due to the fact that the second discussion screens were only completed by 175 students in the computer-based mode, we have focused on the first discussion screen data. Equality of variances was confirmed by a Levene test. Results were not significant, as the ANOVA failed to reveal statistically reliable differences among the three game phases, $F_{(2,713)}=1.016$; $p=0.363$: Individual performance (M=4.15; SD=1.33) is not significantly lower than the correction phase (M=4.30; SD=1.52); and the discussion performance (M=4.34; SD=1.60) is not significantly higher in average than the correction or the individual phases (see figure 5).
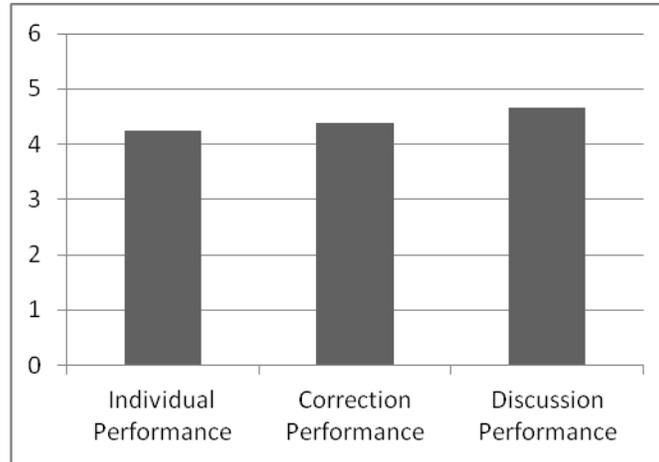

Figure 5. Average performance for Individual and collaborative phases, for the 16 groups of the sample

Hypothesis 2 affirmed that performance would be higher in contexts where students could make explicit and share their LC than those without LC elicitation. In order to measure this hypothesis, we have limited our data to the N=108 students in the six contexts playing with the two different LC options.

H2a: For the individual phase, a two independent samples t-test failed to reveal a statistically reliable difference between individuals elicitating LC (M=4.82; SD=0.933) and individuals playing without the use of the LC tool (M=4.54; SD=0.967), $t(108) =1.542$, $p=0.126$.
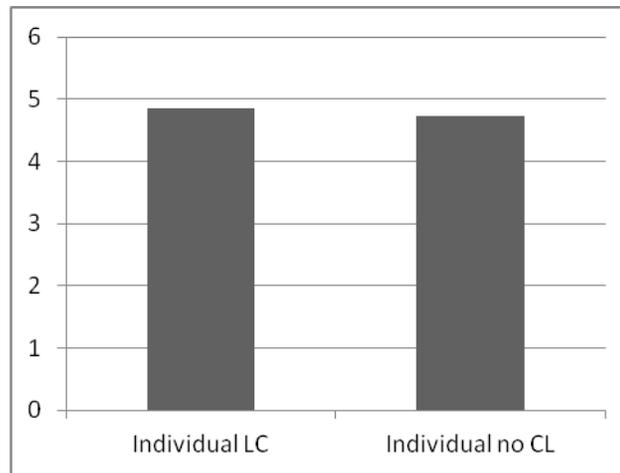

Figure 6. Average scores for the 6 contexts' individual phases with and without LC elicitation

H2b: Collaborative phase performance is also higher for dyads sharing their LC than for dyads without the metacognitive tool. Nevertheless, these differences are not significant. In particular: Correction phase with LC elicitation (M=5.05; SD=1.08) is higher in average than correction phase in

the control group (M=4.75; SD=0.978), t(108) =1.517, p=0.132. Discussion phase had a t(101) =0.776, p=0.434; with LC sharing (M=4.90; SD=1.41) not significantly different from the discussion phase without LC sharing (M=4.67; SD=1.55), we can see the average scorings in figure 7:
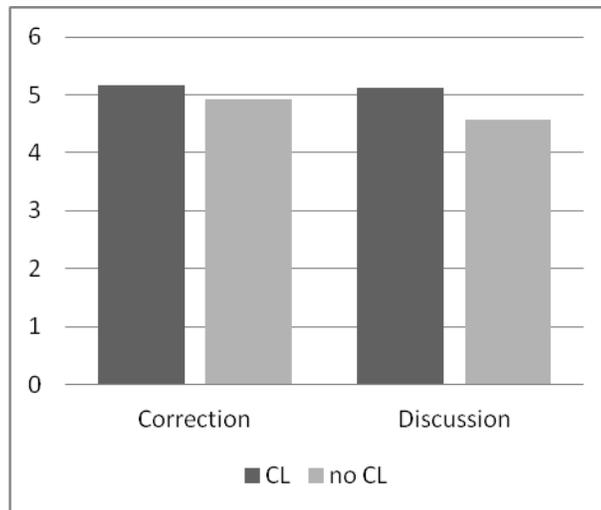

Figure 7. Average collaborative scores for the 6 contexts with and without LC elicitation

7. Discussion

From a first, general analysis, results of the different contexts point to the fact that in all the cases where MetaVals is played as a finance game in HE contexts (ESADE, Carol I), results on performance are higher in average than those experiences with MetaVals in other contexts. In particular, for GaLA and Online Educa conferences, results are lower. This can be related to the fact that participants in these contexts are not students enrolled in a formal course but conference participants playing in a context with higher number of distractions that could be one of the reasons why their results are lower than the students playing in a formal education context. Furthermore, for the subjects of Introduction to Statistics and Learning Theories, a lower performance is observed, both for individual and collaborative phases. Average data show that items in these subjects could be in fact more difficult than those related to finance. However, these differences deserve further study in order to evaluate their significance, and will be object of future analysis with greater samples. The lack of statistical differences could be due to the fact that level of difficulty of the questions is perceived by the participants as low, leading to high number of correct answers. Further research studies involving the MetaVals SG should introduce a higher number of difficult questions and take into account prior knowledge of the participants.

Results for the first hypotheses (H1) show a higher performance in the two collaborative phases, when compared to the individual screen. This improvement in performance for collaborative environments was predicted by Romero and colleagues [21] as results of both intragroup collaboration and intergroup competition processes, purposely designed and implemented in MetaVals SG. However, Hypothesis 1 failed to reveal a significant difference between collaborative and individual performance. This is consistent with previous results on CSCL. In particular, following Davis and colleagues [19], CSCL effectiveness can be influenced by external factors, such as the learners' prior experience, background, availability and expectations, the capacity of technology and technology-based delivery and other resources. Furthermore, as Leemkuil and colleagues [5] admit, some SGs can lack effectiveness when no instructional measures or support are added to guide the learning process. In our case study, although all the tools and virtual guides, most contexts were totally online and asynchronous, this could clearly influence the results in this direction, and hamper the significance of the differences in performance.

Therefore, further versions of the SG need to continue on the adaptation of MetaVals to online, asynchronous settings, most of all when virtual partners' options are played. Following Hess and Gunter [1], the combination of online educational environments and videogames deserves still more research on its learning effectiveness. In this direction, players' prior experience and background data must be studied as possible confounding variables of study hampering the significance of the results. Furthermore, we have seen that collaborative learning itself does not assure better learning results just because two or more students interact [19]; this is in accord with the results, as online interaction is not leading to significant differences in performance when compared to individual achievement. We can understand that there is a need to share different levels of knowledge and metacognitive processes. These leads to the second hypothesis study.

The second outcome variable of the present study, LC, was also measured for all the contexts in relation to game performance, both for individual (H2a) and collaborative phases of the game (H2b). As results show, performance is higher, albeit not significantly, for learners playing with LC elicitation when compared to control group players, both for individual and collaborative phases. The fact that individual performance is higher when students make their LC explicit agrees with the studies on LC elicitation and metacognitive processes: (meta)cognitive judgements do not occur spontaneously in all the learners during the learning process [32]; if there is a guide such our LC tool that helps learners' in the reflection of their own answers, results are better, as learners are aware of their knowledge and their peers' knowledge during the gameplay. Furthermore, following Swartz [33], in multiple choice tasks such as MetaVals SG, the LC elicitation leverages the simplicity of the game, and thus helps players to have a deeper and more significant learning. Nevertheless, the fact that differences between LC and control groups were not significant leads us to think that we need further studies on this hypothesis. In particular, students with similar prior knowledge and in one subject should be studied, as these variables can be predictors of a higher performance [38][39].The similarity or differences of prior knowledge should be further analysed in order to better characterise the knowledge convergence process in SG.

Finally, H2b focused on collaborative contexts. Both correction and discussion phases in MetaVals experiences failed to show significant differences; although performance was higher for LC groups, as predicted by Garner [11]. Nevertheless, the non-significant results could be due to the fact that the vast majority of players engaged in finance contexts, and these gameplays did not show a wide range of performance (students only scored from 4 to 6 in each phase, that is, all scores were high). We should mention that the contexts with lower average performance, which are playing MetaVals version different than financial (see table 1), have a higher degree of difficulty and, therefore, we admit that this question deserves further study in other contexts such statistics (concerning the UWS context) or educational theories (implemented in the UOC cases), where results on performance are more heterogeneous and may allow a deeper analysis without the need of great sample sizes. Furthermore, concerning the study of LC elicitation, students for 2012 and 2013 editions of the Carol I course using MetaVals always showed a (possibly significant) higher LC than the other contexts, the significance of these results may be further studied, focusing especially on the cultural differences, as these students are military, western-Europe profiles different from non-military, eastern profiles in all the other 14 contexts of study. This result deserves further research comparing eastern and western European adult players.

8. Conclusions

Finally, it must be said that researchers decided to maintain LC elicitation and sharing for all the contexts thereafter, as it was observed to help students to have the choice of being aware of their level of knowledge, and thus show better performance [13]. Results found in this 3-year study could set the groundwork for future research in the field of performance and LC in collaborative GBL; in particular, we aim for the implementation of metacognitive tools that could enhance students' sharing and individual reflection of their knowledge and performance during the gameplay. Results also point to the importance of designing collaborative GBL tasks in adult, formal learning courses, such the contexts of study of the MetaVals SG, where learners were eager to be engaged in a GBL activity, and play with another dyad in a computer-based environment. A guide within the game is needed, most of all when MetaVals is played online and asynchronously. In these contexts the second part of the

discussion phase related to the dyads' participation has mostly been ignored by players in the sample. The discussion is not compulsory, which could explain the fact that most of the dyads with virtual peers ignore this part of the game. Therefore, SG in general, and MetaVals in particular has to adapt to each particular context, not only for individual or collaborative gameplays, but also to the specific learners' needs such contents level, communication within the game, and modality. The results of this 3 years study developed in the context of MetaVals points to the interest of collaborative learning in GBL activities and the need for supporting metacognitive judgements, such the LC, in GBL. Nevertheless, further studies using other similar SGs should be conducted in order to ensure the generalisation of the results in other collaborative game activities.

9. References

[1] Hess, T., Gunter, G., "Serious game-based and nongame-based online courses: Learning experiences and outcomes" British Journal of Educational Technology, vol. 44, no. 3, pp. 372–385, 2013. http://dx.doi.org/10.1111/bjet.12024

[2] Bellotti, F., Berta, R., De Gloria, A., Lavagnino, E., Dagnino, F., Ott, M., Romero, M., Usart, M., Mayer, I., "Designing a Course for Stimulating Entrepreneurship in Higher Education through Serious Games." Procedia Computer Science, vol. 15, pp. 174-186, 2012. http://dx.doi.org/10.1016/j.procs.2012.10.069

[3] Mawdesley, M., Long, G., Al-Jibouri, S., Scott, D., "The enhancement of simulation based learning exercises through formalised reflection, focus groups and group presentation", Computers and Education, vol. 56, no. 1, pp. 44-52, 2011. http://dx.doi.org/10.1016/j.compedu.2010.05.005

[4] Dickey, M. D., "Game Design and Learning: A Conjectural Analysis of How Massively multiple Online Role-Playing Games (MMORPGs) Foster Intrinsic Motivation", Educational Technology Research and Development, vol. 55, no. 3, pp- 253-273, 2007.

[5] Leemkuil, H., de Jong, T., de Hoog, R., Christoph, N., "KM Quest: A Collaborative Internet Based Simulation Game", Simulation & Gaming, vol. 34, no. 1, pp. 89-111, 2003. http://dx.doi.org/10.1177/1046878102250605

[6] Herz, C. J., "Gaming the system: What higher education can learn from multiplayer online worlds. The Internet and the University", Educause Forum on the Future of Higher Education. Retrieved [11/06/2013] from http://www.educause.edu/ir/library/pdf/ffpiu019.pdf

[7] Azevedo, R., Witherspoon, A., Chauncey, A., Burkett, C., Fike, A., "MetaTutor: A MetaCognitive tool for enhancing self-regulated learning". In R. Pirrone, R. Azevedo, & G. Biswas (Eds.), Proceedings of the AAAI Fall Symposium on Cognitive and Metacognitive Educational Systems (pp. 14-19). Menlo Park, CA: Association for the Advancement of Artificial Intelligence (AAAI) Press, 2010

[8] Azevedo, R., "The role of self-regulation in learning about science with hypermedia". In D. Robinson & G. Schraw (Eds.), Recent innovations in educational technology that facilitate student learning. (pp.127–156), 2010

[9] Järvelä, S., Volet, S., "Motivation in Real-Life, Dynamic, and Interactive Learning Environments: Stretching Constructs and Methodologies". Journal European Psychologist, vol. 9, no. 4, pp. 193-197, 2004. http://dx.doi.org/10.1027/1016-9040.9.4.193

[10] Merenluoto, K., Lehtinen, E., "Number concept and conceptual change: towards a systemic model of the processes of change", Learning and Instruction, vol. 14, no. 5, pp. 519-534, 2004. http://dx.doi.org/10.1016/j.learninstruc.2004.06.016

[11] Garner, R., "When children and adults do not use learning strategies: Toward a theory of settings", Review of Educational Research, vol. 60, pp. 517–529, 1990 http://dx.doi.org/10.3102/00346543060004517

[12] Schraw, G., "Promoting general metacognitive awareness", Instructional Science, vol. 26, pp. 113–125, 1998 http://dx.doi.org/10.1023/A:1003044231033

[13] Romero, M., Usart, M., Popescu, M., Boyle, E. "Interdisciplinary and International Adaption and Personalization of the MetaVals Serious Games", Lecture Notes in Computer Science, vol. 7528, pp 59-70, 2012. http://dx.doi.org/10.1007/978-3-642-33687-4_5

[14] Peirce, N., Conlan, O., Wade, V., "Adaptive educational games: Providing non-invasive personalised learning experiences". In Digital Games and Intelligent Toys Based Education, 2008 Second IEEE International Conference on IEEE, pp. 28-35, 2008 http://dx.doi.org/10.1109/digitel.2008.30

[15] Sawyer, B., "The serious games summit: emergent use of interactive games for solving problems is serious effort", Computers in Entertainment (CIE) - Theoretical and Practical Computer Applications in Entertainment, vol. 2, no. 1, pp. 5-12, 2004. http://dx.doi.org/10.1145/973801.973811

[16] Moreno-Ger, P., Burgos, D., Sierra, J.L. & Fernández-Manjón, B., "Educational Game Design for Online Education", Computers in Human Behavior, vol. 24, no. 6, pp. 2530-2540, 2008 http://dx.doi.org/10.1016/j.chb.2008.03.012

[17] Dillenbourg, P., Jarvela, S., Fischer, F., "The evolution of research on computer-supported collaborative learning: From design to orchestration", Technology-Enhanced Learning, vol. 1, pp. 3–19, 2009 http://dx.doi.org/10.1007/978-1-4020-9827-7_1

[18] Wendel, V., Gutjahr, M., Göbel S., Steinmetz, R., "Designing collaborative multiplayer serious games", Education and Information Technologies, vol. 18, no. 2, pp. 287-308, 2013 http://dx.doi.org/10.1007/s10639-012-9244-6

[19] Lipponen, L., Rahikainen, M., Lallimo, J., Hakkarainen, K., "Patterns of participation and discourse in elementary students' computer-supported collaborative learning", Learning and Instruction, vol. 13, pp. 487–509, 2003 http://dx.doi.org/10.1016/S0959-4752(02)00042-7

[20] Kreijns, K., Kirschner, P.A., Jochems, W., "Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research", Computers in Human Behavior, vol. 19, no. 3, pp. 335-353, 2003 http://dx.doi.org/10.1016/S0747-5632(02)00057-2

[21] Davis, A., Little, P., Staward, B., "Developing an infrastructure for online learning", In Anderson, T. (ed.) The theory and practice of online learning. Edmonton: AU Press, 2008

[22] Prensky, M., "The Motivation of Game Play or, the REAL 21st century learning revolution", On The Horizon, vol. 10, no. 1, 2001

[23] Romero, M., Usart, M., Ott, M. ,Earp, J. , de Freitas, S., Arnab, S., "Learning through playing for or against each other? Promoting collaborative learning in digital game based learning", ECIS 2012 Proceedings. Paper 93. Retrieved November 10, 2012, from http://aisel.aisnet.org/ecis2012/93

[24] Becker, K., "Digital game-based learning once removed: teaching teachers. British Journal of Educational Technology, vol. 38, no. 3, pp. 478-488, 2007 http://dx.doi.org/10.1111/j.1467-8535.2007.00711.x

[25] Jeffrey, L., "Learning orientations: Diversity in higher education", Learning and Individual Differences, vol. 19, pp. 195–208, 2009 http://dx.doi.org/10.1016/j.lindif.2008.09.004

[26] Munneke, L., Andriessen, J., Kanselaar, G., Kirschner, P., "Supporting interactive argumentation: Influence of representational tools on discussing a wicked problem", Computers in Human Behavior, 23, no. 3, pp. 1072, 2007 http://dx.doi.org/10.1016/j.chb.2006.10.003

[27] Ducheneaut, N., Yee, N., Nickell, E., Moore, R. J., "Alone Together?" Exploring the Social Dynamics of Massively Multiplayer Online Games (CHI Conference), pp. 407-416, 2006 http://dx.doi.org/10.1145/1124772.1124834

[28] Roebers, C. M., "Confidence judgments in children's and adults' event recall and suggestibility", Developmental Psychology, vol. 38, pp. 1052–1067, 2002 http://dx.doi.org/10.1037/0012-1649.38.6.1052

[29] Kirschner, P.A., Erkens, G., "Cognitive Tools and Mindtools for Collaborative Learning", Journal of Educational Computing, vol. 35, no. 2, pp. 199–209, 2006 http://dx.doi.org/10.2190/R783-230M-0052-G843

[30] Valdez, A., "Student Metacognitive Monitoring: Predicting Test Achievement from Judgment Accuracy", International Journal of Higher Education, vol. 2, no. 2, 141-146, 2013 http://dx.doi.org/10.5430/ijhe.v2n2p141

[31] Karpicke, J. D., Butler, A. C., Roediger III, H. L., "Metacognitive strategies in student learning: do students practise retrieval when they study on their own?", Memory, vol. 17, no. 4, 471-479, 2009 http://dx.doi.org/10.1080/09658210802647009

[32] Swanson, H. L., "Influence of metacognitive knowledge and aptitude on problem solving", Journal of educational psychology, vol. 82, no. 2, pp. 306, 1990 http://dx.doi.org/10.1037/0022-0663.82.2.306

[33] Swartz, S. M., "Acceptance and accuracy of multiple choice, confidence-level, and essay question formats for graduate students", Journal of Education for Business, vol. 81, no. 4, pp. 215-220, 2006 http://dx.doi.org/10.3200/JOEB.81.4.215-220

[34] Flannelly, L., "Using feedback to reduce students' judgment bias on test questions", Journal of Nursing Education, vol. 40, no. 1, 10-16, 2001

[35] Efklides, A., Akilina, S., Petropoulou, M., "Feeling of difficulty: an aspect of monitoring that influences control", European Journal of Psychology of Education, vol. 15, pp. 461–476, 1999 http://dx.doi.org/10.1007/BF03172973

[36] Brennan, S.E., Chen, X., Dickinson, C.A., Neider, M.B., Zelinsky, G.J., "Coordinating Cognition: The Costs and Benefits of Shared Gaze during Collaborative Search", Cognition vol. 106, no. 3, pp. 1465– 1477, 2008 http://dx.doi.org/10.1016/j.cognition.2007.05.012

[37] Bluemink, J., Hämäläinen, R., Manninen, T., Järvelä, S., "Group-level analysis on multiplayer game collaboration: how do the individuals shape the group interaction?" Interactive Learning Environments, vol. 18, no. 4, pp. 365-383, 2010 http://dx.doi.org/10.1080/10494820802602444

[38] O'Reilly, T., McNamara, D.S., "The Impact of Science Knowledge, Reading Skill, and Reading Strategy Knowledge on More Traditional "High-Stakes" Measures of High School Students' Science Achievement", American Educational Research Journal, vol. 44, no. 1, pp. 161-196, 2007 http://dx.doi.org/10.3102/0002831206298171

[39] Hailikari, T., Nevgi, A., Komulainen, E., "Academic self-beliefs and prior knowledge as predictors of student achievement in Mathematics: a structural model", Educational Psychology: An International Journal of Experimental Educational Psychology, vol. 28, no. 1, pp. 59-71, 2008 http://dx.doi.org/10.1080/01443410701413753