Evaluating the Effects of Orchestrated, Game-Based Learning in Virtual Environments for Informal Education

Panagiotis Apostolellis
Virginia Tech
2202 Kraft Drive, Blacksburg, VA 24060
panaga@vt.edu

Doug A. Bowman
Virginia Tech
2202 Kraft Drive, Blacksburg, VA 24060
dbowman@vt.edu

ABSTRACT
In informal learning spaces employing digital content, such as museums, visitors often do not get adequate exposure to content, or they passively receive instruction offered by a museum docent to the whole group. This research aims to identify which elements of co-located group collaboration, virtual environments, and serious games can be leveraged for an enhanced museum learning and entertaining experience. We developed C-OLiVE, an interactive virtual environment supporting tripartite group collaboration, which we used to explore our hypothesis that synchronous, co-located, group collaboration will afford greater learning compared to conventional approaches. In an empirical study, we found some evidence supporting this hypothesis, taking into consideration other factors such as game experience and social presence. Students participating in the three-player condition demonstrated a better understanding of the collaborative tasks compared to their single-player counterparts. We discuss these results and outline future studies using the same virtual environment.

Author Keywords
Collaborative learning; virtual environments; serious games; contextual learning; museums.

ACM Classification Keywords
• **Applied computing ~ Interactive learning environments**
• *Applied computing ~ Collaborative learning* • Human-centered computing ~ Empirical studies in HCI

INTRODUCTION
There has been quite some interest during recent years in bringing the benefits of games and their entertainment value into the education arena. Many researchers are working towards this direction with the objective to exploit the pedagogical benefits of gaming technologies both in formal and informal learning spaces [3,24,32]. Part of the serious games initiative is guided towards this direction as well, with the intention to motivate children by actively engaging them in the learning process. Indeed, this generation of learners, or "digital natives" as they have been characterized [25], demand a significantly different instructional approach than the one offered by traditional pedagogy.

Since reform of formal education is a challenging and time-consuming process, it makes sense to start by incorporating these novel educational practices in informal learning. A significant requirement for this reform is that "learning can and must become a daylong and lifelong experience," occurring "not only in schools, but also in homes, community centers, museums, and workplaces" [26].

Additionally, researchers have argued that the learning approaches used must depart from objectivism (i.e., knowledge existing independent of learners' minds) towards constructivism, where learning occurs through socially active learning environments, through the interactions of people and artifacts [36]. Previous research has shown that virtual environments are a fruitful arena for constructivist learning [10,28,35], but rarely have these practices been integrated with the entertaining character of museum visits.

Our research is focused on exploring the opportunities afforded by social experiences in museum spaces combined with serious games and collaborative virtual environments (CVEs). In our first study, described in this paper, we built a VE for multi-user interaction and tested it in different collaborative conditions. We have exploited the co-located collaborative nature of museum visits to enhance the system with important features of CVEs, like shared context, awareness of others, and negotiation and communication [31]. Our hypothesis was that greater levels of interaction and collaboration would afford increased engagement and learning. Besides learning gains, we assessed the motivational impact of social presence, game experience, and presence (i.e., the sense of being inside the virtual world) on learning.

RELATED WORK
Learning in VEs has been studied for almost two decades—since the time when virtual reality technology became accessible to research labs. Perhaps the most influential work so far is the *ScienceSpace* project where researchers assessed the impact of VE properties like multisensory immersion, multiple representations, and interaction for teaching complex and abstract scientific concepts [9]. These re-

ACE '14, November 11 - 14 2014, Funchal, Portugal

http://dx.doi.org/10.1145/2663806.2663821

searchers found some attributes of the technology promising as a constructivist tool for learning science. Work that followed involved the use of collaboration inside VEs as a means to increase engagement and learning through peer activities. In some cases, elementary school students worked either in virtual dyads wearing head-mounted displays [18], or in teams from remote locations using different multiuser VE technologies [28]. Although no specific learning gains could be attributed to collaboration, researchers of the *NICE* project [28] reported that this might be due to inadequate instructional design in order to support the learning activities through cooperative work.

At the same time, researchers were starting to value the power of games and simulations for facilitating learning through interactive learning environments, especially for children [19,27,32]. These applications are "an integral part of their social and cultural lives" and children "often evaluate their status in a peer group based on their interaction in games" [27]. Moreover, besides being fun, games can be intrinsically motivating; i.e., the goal of the game is the same as the learning objective. Significant motivating factors of a game that can facilitate learning, as defined by Malone and Lepper [23], are fantasy, curiosity, challenge, and control, whereas problem-solving, logic, memory, and visualization are some of the skills perceived to be important in the popular genre of adventure games [2].

Simulations have also been widely used for training and learning purposes. Their main distinction from games is that they represent some type of realistic setting or situation which would be difficult to experience in real life (due to cost, scale, location, or other constraints). Learners using simulations have been shown to better understand the relationship of the content with real-life experiences upon reflection, since the simulation is closely connected with the learning context [15]. A combination of games and simulations can be found under the label of serious games, which has gained a lot of attention in research and practice recently. The work of Bellotti et al. [4] is just one case where multi-user serious games are used for cultural heritage applications, employing a task-based learning approach by breaking down knowledge constructs in mini games.

In cases like this, collaboration among participants is used both as a means to facilitate learning through communication of intentions, negotiation of actions, and reflection of outcomes, as well as a motivating factor by itself. Such work falls usually under the title of computer-supported collaborative learning (CSCL) and has been extensively studied. Popular recent CSCL applications include the genres of multi-user virtual environments (MUVEs) and massively multiuser online games (MMOGs). Although not explicitly intended for learning, MUVEs and MMOGs, such as *Second Life*, *Action Worlds*, and *World of Warcraft*, have been frequently used for educational purposes inside [33] and outside [34] the classroom. In most cases, researchers report increased motivation and engagement deriving from the gaming and immersive attributes of these systems, eventually leading to enhanced learning outcomes.

Besides using commercial online MUVEs, researchers have also experimented with custom-made virtual worlds for learning. *Quest Atlantis* is a 3D multi-user environment used by thousands of students within classrooms for intentional learning [3]. Students immersed in the game as protagonist learn through transformational play, which involves employing conceptual understanding of information in order to change a problem-based fictional context. Comparing a game-based and a story-based lesson on persuasive writing revealed that the game-based platform afforded significantly greater learning gains and engagement levels. *River City* is another custom MUVE where students learn about hypothesis formation and experimental design through collaboratively identifying and making inferences about problems in epidemiology [11]. Students' use of the VE simulation of a 19^{th}-century city resulted in substantial gains in knowledge and skills in scientific inquiry, as compared to conventional instruction or an equivalent board game.

These MUVEs, however, are distributed and are missing the benefits of co-located collaboration deriving from the physical co-presence of other learners. The main attribute affected by co-located collaboration is social presence, which is the sense of being together with someone else. Studies have shown that playing together with or against others while in the same space yields greater enjoyment, challenge, and perceived competence [e.g., 16]. It would be interesting then, to investigate the effects of this positive influence on learning. In another study, access to nonverbal cues through video communication while in a networked CVE did not result in increased social presence as compared to just audio, which in part was attributed to the use of avatars [29]. Furthermore, the mere co-location of collaborators does not always imply greater social presence, since other factors such as physical environment and application design affect communication [21].

Thus, further research is needed on the benefits of co-located interaction within a group of students (i.e., more than two), especially in the socially rich learning spaces of museums. We need to understand how the transparency of decision outcomes in co-located collaboration (i.e., feeling/seeing the other persons' reactions to your choices) affects the decision-making process [20], and consequently the learning achieved. Museum spaces, in particular, provide ample opportunities for learning through informal social interactions between visitors, digital exhibits, and museum educators. We believe that the educational potential that emerges from the convergence of serious games with other technologies, like virtual environments, must be supported by appropriate pedagogical practices, peer interactions, and tutor (or facilitator in case of a museum) engagement [14]. This identified need led us to investigate the unexplored space that lies in the convergence of serious games, co-located CVEs, and informal education.

Figure 1. The C-OLiVE virtual olive oil factory used as a test bed for assessing our hypotheses.

C-OLIVE: COLLABORATIVE ORCHESTRATED LEARNING IN VIRTUAL ENVIROMENTS

In order to explore our research questions, we developed a virtual environment that supports co-located collaboration of up to three players.

CVE Design

C-OLiVE (Collaborative Orchestrated Learning in Virtual Environments) is an accurate representation of an actual steam-powered olive oil factory of the 1950s from the island of Lesvos in Greece, containing exact 3D replicas of the machinery used at the time (see Figure 1). The objective of the game is to learn about olive oil production. Initially, the players have to produce steam in order to start up the olive oil production machinery, which they then have to operate. The operation involves fifty different tasks that need to be executed in the factory, using eight items/tools, and interacting with thirty different machinery/process parts. Many tasks are recurring to provide ample opportunities for interaction. The players are led through the processes (either steam or olive oil production) through the alerts that appear at the bottom of the map or their own window, if they are close enough to the workstation involved. They can move freely inside the virtual factory and have to visit the workstations in order to identify and troubleshoot each problem on a timely manner. There are twelve different types of workstations (i.e., factory machinery like a steam engine, a boiler, mills, presses, and pumps) that players have to operate either individually or collaboratively. There are five distinct collaborative tasks, which two players have to execute synchronously by applying a tool on the same or different machine parts (e.g., in Figure 1 Doug used a wrench to help Panagiotis attach a belt on a machine). Most such tasks are repetitive to allow a different pair of players to execute them each time (there are thirteen collaborative tasks in total). When only one player is controlling the game, an icon indicates that the task is collaborative and a second trigger within five seconds executes the action.

The VE was built using the Vizard VR toolkit, and 3D Studio Max and Maya for the construction of the 3D models. The system is designed to support one to three simultaneous players and different kinds of game controllers; we have decided to use wireless Xbox controllers due to their popularity as gaming devices and rich available inputs. A single large front-projected display (non-stereoscopic) is used by all players with individual viewpoints. Our initial design involved a combination of individual and shared viewpoints depending on the task at hand, but we eventually decided that this might confuse the players, in addition to decreasing autonomy as a result of the restricted navigation. Control and pace has been shown to be a catalyst for effective assimilation of information in museums and similar spaces of exploratory learning [1,13]. Thus, each player is provided with a dedicated viewpoint in a separate window.

Each window has a side panel containing vital information about the game. The player's name and color, her inventory, and feedback for any action attempted are all displayed there. When the player selects a tool from the inventory it is highlighted and appears in the middle of her viewport (e.g., the bottom-right player in Figure 1 has selected the wrench). A top-down map of the factory is also shown at the top-right of the display; it contains avatar locations, warnings and information messages about the status of the machinery, and the game score. Players have to attend to the alerts on the map and troubleshoot problems as quickly as possible. Points are added for every correct action while points are lost when a problem is not solved for a prolonged period.

A finite state machine (FSM) is used to handle all the events happening in the factory based on user input and the current state of each machine. An FSM has been defined for each machine, and a factory FSM monitors all events and distributes messages and actions to the machine(s) involved in an interaction (e.g., a player applies a tool on a machine part) or an event (e.g., a timer expires indicating that a process finished). Moreover, the machinery FSM contains all the information, warning, and interaction feedback messages that are presented during game play.

Situated Orchestrated Learning

The basic premise of learning occurring in VEs is the same as the one for free-choice learning spaces like museums. This has to do with the fact that the protagonist of the fictional world is engaged in realistic activities that can change the fate of the world [3], in a similar fashion as the visitor of a museum is engaged with authentic activities that reconstruct real-life experiences and concepts [13]. Similarly, learning in C-OLiVE is achieved by engaging the students in authentic activities within an actual olive oil factory and having them negotiate their actions in the physical space, afforded by the co-located collaborative nature of the game. Situating learners in authentic problem-solving activities is believed to be one of the major benefits of virtual environments [11].

We use the term *orchestrated learning* to denote the type of learning that occurs through the intentional facilitation of coordinated collaborative activities within the VE. Others have used the same term to describe the "coordination of learning episodes," as a means of orchestrating the interactions between learners and resources within a classroom [7]. As an example, in C-OLiVE collaboration is demanded in different parts of the game where workers in a real factory had to complete a task by working together (e.g., lift the heavy sacks full of olives). When a collaborative action is executed the players involved are notified with a sound and an icon; in Figure 1 the top-left and bottom-right players have completed such a task. Encouraging collaborative activities by design through the demand of inter-group coordination increases social (inter)actions between participants [6] and is believed to have a positive impact on learning.

Adopting this metaphor of learners as members of a musical orchestra has other connotations for our instructional approach besides emphasizing the demand and benefits of coordinated action. Musicians are totally engaged in their performance and their actions are driven by their intrinsic passion for excellent acoustic results, similar to how games like C-OLiVE match the goal and process of the game with the learning outcomes. This engagement with free-choice tasks where extrinsic rewards are absent has been connected to increased sense of the *flow experience* [8], but also to enhanced learning due to its intrinsically motivating power [23]. Additionally, the increased levels of fun and engagement deriving from peer collaboration within groups visiting museums [13] or working in classrooms [15], might in some cases reinforce learning outcomes. Finally, similar to how musicians participate in a shared community of practice where common goals and understanding are necessary for its success [22], C-OLiVE provides a (virtual) community of workers where players have to negotiate their actions and plan a common route through distributed decision-making.

STUDY

The main goal of this study was to understand the impact of interactivity on learning in collaborative game play. We wanted to move beyond the typical two-player setup and incorporate a third player, in order to increase the resemblance of the setting to group interaction, which is common in museum visits. The two main research questions we were trying to address have to do with interactivity and the social presence that derives from the co-located aspect of C-OLiVE. More specifically, these are:

RQ1: What is the effect of the level of interactivity on learning in a gaming CVE?

RQ2: How is learning affected by the within-group collaboration in physical space?

By consolidating the results of prior work and our own insights from working with children in museums, we created a path model with all the interactions we believe are important in our study (see dashed rectangle in). The lines indicate relationships between the constructs and the plus (+) sign indicates a positive correlation; e.g., we hypothesize that increased interactivity will positively affect engagement, presence, social presence, and learning. Based on this model we designed the experiment described below.

Experimental Design

Since interactivity was our main interest in this study, we used it as our between-subjects independent variable and designed three different levels. These had to do with the degree of involvement of the participants in the game and are the following: *auto*, where all the learners are watching a recording of someone playing the game; one-player (*1P*), where one player was controlling the game and the rest were helping her by indicating problems that needed attention or suggesting plans of action; and three-player (*3P*), where all three players were interacting directly with the game. The first setup is typical of existing museum guides for groups, where a curator or museum educator is controlling a digital exhibit, but we opted for a more controlled setup with no human intervention. In some cases one of the participants gets to acquire control while the rest are indirectly participating, which resembles our second experimental condition. Finally, the third condition is the one we hypothesize will afford greater learning benefits, due to increased engagement and social presence.

The main dependent variable was learning, while game experience and social presence were the secondary ones. Hence, we decided to administer a pre-test and a post-test and measure learning as the difference between the two scores. The test was a quiz about the content of the game (i.e., olive oil production in a steam-powered factory of the mid-19th century in Greece) and was divided into three parts. The first part (12 questions) asked general questions about factual information presented in the game; the second part (5 questions) assessed higher-level knowledge of the domain and demands that students combine information presented in the game and extrapolate their responses; and the last part (10 questions) evaluated their understanding of collaborative tasks performed in the factory[1].

To measure game experience, as a superset of engagement, we used the game experience questionnaire (GEQ) developed by the Game Experience Research Lab at the University of Eindhoven [17]. We found this an appropriate self-report instrument for measuring game experience, as it was more sensitive to a broader range of emotions associated with digital game experiences and also provided a version specifically for kids. The questionnaire is broken down into seven different dimensions that contribute, positively or negatively, to the game experience: *immersion* in the story (closely related to engagement; not to be confused with the attribute of the display technology), *flow* [8], *challenge*, *competence*, *positive* and *negative feelings*, and *tension*. Although not specifically designed for learning games, we believe this multi-measure approach can capture the different emotional states that a game player goes through.

A complementary questionnaire developed by the same research group, the social presence in gaming questionnaire (SPGQ) [21], was used to capture the effect of the social context and interactions on the digital game experience. Although other instruments were considered [5], the SPGQ was chosen mainly because of its appropriateness for evaluating social presence in gaming settings. This distinction arises from the continuous interdependence noticed in gaming VEs, which is not usually the case in teleconferencing systems [21]. Additionally, the three constructs of the SPGQ were appropriate for the co-located nature of the game: psychological involvement was described by positive (*empathy*) and *negative emotions* towards co-players, while *behavioral involvement* captured the degree of interdependence, which is significant in game experiences.

We used three questions for each one of the seven GEQ measures and a total of twelve questions from the SPGQ, all taken from the child versions of the questionnaires. Finally, we used the Slater, Usoh, Steed (SUS) questionnaire for assessing presence [30] (i.e., the feeling of 'being' inside the virtual factory). Although presence was not included in our research questions and our setup was not fully immersive, we still thought it would be interesting to measure presence to evaluate its effects on learning.

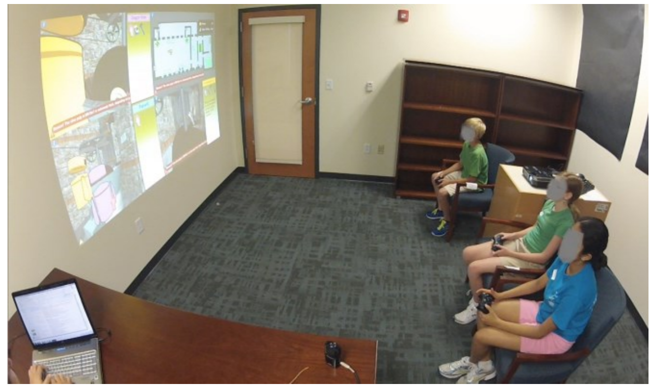

Figure 2. Experimental setup (controlled) of the 3P condition.

Participants

Our goal was to have a sample size of 54 students, based on an *a priori* power analysis (80%) for a medium effect of interaction level on learning. We eventually conducted a controlled study with 47 middle school students between 11 and 14 years old (M=12.25). We had more than twice as many males than females (33 boys, 14 girls). The distribution among conditions, which was randomly made based on pre-assigned groups, was the following: five groups (14 students) for the auto condition (one student did not show up), six groups (18 students) for the 1P condition, and five groups (15 students) for the 3P condition. Eventually, the study was conducted during a period of almost six months and students were recruited from two summer camps and two private schools from the area near the university.

Apparatus

The game was projected on a large front-projected 16:9 wall display, and students were sitting at a distance of around 10ft (see). A PC with an Intel Core i7 Extreme processor and a GeForce 7950 GT graphics card was used to run the application; we chose such a powerful setup in order to cope with the processing demands of rendering the environment three times, one for each individual viewpoint, in the 3P condition. Wireless Xbox controllers were used as interaction devices. The post-study questionnaires were administered online using iPads to enable automatic recording of responses.

Procedure

Student participants were approached through the managing entity, either the summer camp director or the class teacher, and parental permission was granted. The background survey and pre-test were administered online fifteen days prior to the experiment to avoid priming. Participants were instructed through a letter to their parents to complete the quiz to the best of their knowledge without any assistance from any source. During completion of all forms, students had to select a code name, which was used to identify their responses and match them with the game data, but also their real names in case we wanted to follow up on the study. On the day of the experiment, groups of three students were

[1] The study material can be accessed from the following url: http://bit.ly/c-olive

initially informed about the goal of the study and watched a five-minute video explaining the main processes involved in steam and olive oil production. Next, they participated in a practice session to get used to the controls and the game interface. As soon as everyone felt comfortable (no time limit was imposed) they started the main trial.

Besides video recording each session (after obtaining verbal assent), a variety of time-stamped game data were recorded: actions attempted (errors) and performed, stations visited in the factory (using enter and exit triggers), full location log (every 0.5 seconds), machine state changes (both due to user interaction and timer events), score, and game duration. After the game was completed (again, no time limit was imposed and there was no way to lose), participants took the presence and game experience questionnaires, and then took the post-test (which was exactly the same as the pre-test). This order of survey administration was used to avoid recall of information from short-term memory. Participants in the auto condition, who watched a video of someone playing the game, did not complete the game experience or social presence questionnaires; neither did they have any game data being logged. Last, we conducted a short interview about participants' experiences with the game.

RESULTS

Initially, we examined some exploratory statistics and graphs in order to find extreme values that might skew the data. We found two cases where participants gave the highest score on all game experience measures, apparently not noticing that there were also negative measures included (e.g., tension). Any value that was more than two standard deviations higher than the mean was considered an outlier and was substituted with the next highest value (also known as Winsorizing). Throughout our analysis, we made sure that the statistical models used satisfied the necessary assumptions and corrected values are reported whenever these assumptions were violated.

Although participants were randomly assigned to the three conditions by the summer camp or school administrator, we wanted to eliminate any possible confounds that might arise due to uncontrolled factors. Thus, we first explored Pearson correlations and then ran a series of analyses of covariance (ANCOVA) using our demographic and other data gathered from the background survey as covariates. Besides age, gender, and school grade we also asked about exposure to computer games, MUVEs, VR technology, collaborative games, and domain knowledge. None of these data were shown to affect learning. We only noticed a significant positive correlation between *hours playing games per day* and learning ($F(1,43) = 5.87$, $p = 0.02$, $r = 0.35$); however, there was no significant effect of interaction level on learning after controlling for the effect of this covariate ($F(2,43) = 0.98$, $p = 0.38$, partial $\eta^2 = 0.044$).

Despite our attempts to ensure that participants take the pre-test well in advance of the experiment, it was impossible to control this factor. It was observed that students took the test in many cases on the day just before the experiment, and in nine cases on the day of the study. Consequently, we ran an ANCOVA with *days prior to the experiment* as the covariate, and found no effect of the time between tests on learning.

Effect of Interaction Level on Learning

In order to analyze the effect of our independent variable on learning, we conducted a mixed one-way ANOVA, where the time the test was taken was the within-subjects variable (pre- vs. post-test) and interaction level was the between-subjects variable (Auto vs. 1P vs. 3P). Participants in all three conditions revealed a highly significant gain in learning both overall and on each individual part of the test ($p < 0.001$), as can be seen in (all bars show std. error).

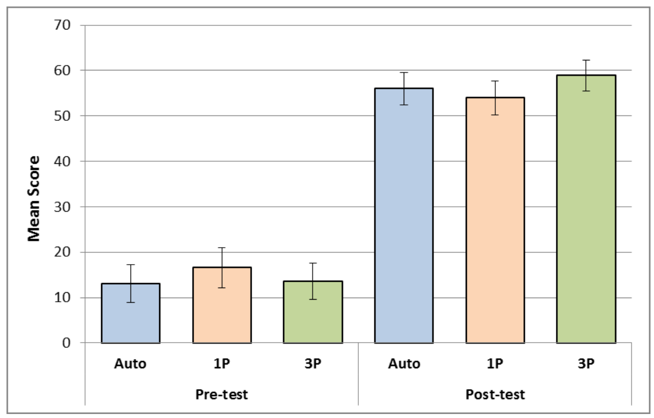

Figure 3. Effects of interaction level on learning.

However, there were no significant differences between the three conditions on the overall learning score. There was a significant effect of interaction level on learning for the last part of the quiz about collaborative tasks ($F(2,44) = 3.465$, $p = 0.04$). A Games-Howell multiple comparisons test revealed that students in the 3P condition demonstrated more learning on this section than the 1P ones ($p = 0.036$). This effect is diminished when only the non-controllers are taken into account for the 1P condition ($F(2,38) = 2.82$, $p = 0.072$), indicating that the controllers in the 1P condition had a poorer understanding of the collaborative tasks. This counterintuitive finding might be attributed to the added cognitive load needed for controlling the game.

Although the level of control of the game was not an explicitly controlled independent variable in our study, we still deemed it interesting to investigate how the level of control predicts learning. Running an ANOVA with time of test as the repeated measures variable and degree of control as the between-subjects variable with three levels (no control, controlling alone, controlling with others) we found again no significant effects on learning, except for the final collaborative part, where score difference was significantly affected by control level ($F(2,44) = 3.529$, $p = 0.038$, $r = 0.27$). A post hoc Games-Howell test revealed that the difference was due to the players in the 3P condition who performed better than the ones who had no control whatsoever, either in the Auto or 1P conditions ($p = 0.044$). Although this appears to be a medium effect ($r \sim 0.30$), we should mention that the largely unequal sample sizes are a matter of concern for the reliability of post hoc tests, especially with the small sample size of 1P controllers ($N = 6$).

We also ran individual independent-samples t-tests for the 1P condition ($N = 18$) between the controllers ($N = 6$) and non-controllers ($N = 12$) and found no significant differences in learning. Additionally, we ran one-way ANOVA and post-hoc tests on each individual test question for a more fine-grained investigation of the effect of our independent variable on learning. We did not find any patterns that can help us draw useful conclusions, besides verifying that the 3P-condition participants had a better understanding of specific collaboration-related questions, especially when compared to the 1P ones. A surprising finding was the significantly better learning of the Auto-condition participants on one of the higher-level questions (i.e., *"What might have caused the press tray to break?"*) as compared to the 3P participants. In the discussion section below, we speculate on why the Auto condition might prove beneficial in some cases.

Effects of Interaction Level on Game Experience

Scores on the GEQ and SPGQ were measured only for the interactive conditions on a scale of 1 (not at all) to 5 (extremely). Running a multivariate ANOVA on all seven measures, we found a significant effect of interaction level on GE using Pillai's trace ($V = 0.47$, $F(7,25) = 3.13$, $p = 0.016$), which indicates that students in the 3P condition ($N = 15$, $M = 4.18$, $SD = 0.26$) enjoyed the game much more than their 1P counterparts ($N = 18$, $M = 3.78$, $SD = 0.65$). Also, running individual ANOVAs for these measures (), we found a significant effect of interaction level on both positive ($p < 0.05$) and negative affect ($p < 0.001$), which reveals how the degree of interactivity affected the participants' emotional state. For any GE measures where the assumption of the homogeneity of variance was violated, a Brown-Forsythe corrected F-statistic was used.

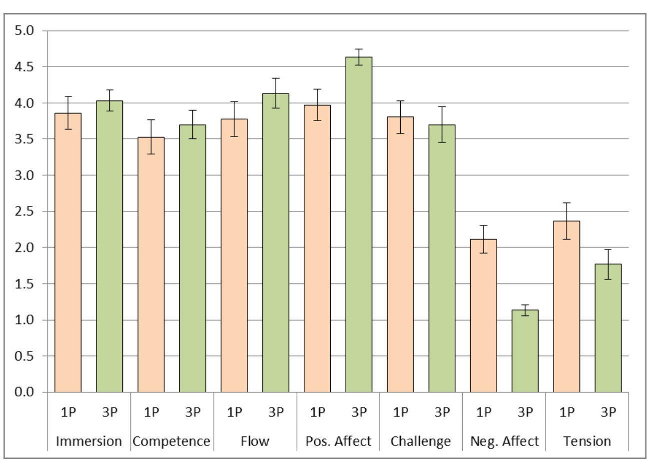

Figure 4. Results for the individual GE measures.

None of the remaining game experience constructs was significantly affected by the treatment. Finally, there was no difference in presence score among the three conditions.

Effects of Interaction Level on Social Presence

Similarly to the GE analysis, we first ran a MANOVA with all three measures of SP, but found no significant effect of interaction level on social presence. However, the individual ANOVAs for each measure indicated significantly increased negative feelings for the 1P condition (Brown-Forsythe corrected $F = 5.58$, $p < 0.026$). This effect was diminished when only the non-controllers were taken into account, which is an interesting finding considering that we would expect these participants to be less satisfied than the ones controlling the game. To investigate this finding further, we compared the means of the participants split by the level of control (1P control or 3P control) and no control at all (i.e., the 1P non-controllers). The results shown in Figure 5 reveal an apparent difference both in behavioral involvement and negative feelings. Although not statistically significant (only the uncorrected statistic of negative feelings is marginally significant at $p = 0.046$), the post hoc tests indicate significant differences between some of the three conditions. Specifically, a Tukey HSD test revealed a significant increase of behavioral involvement from 1P no control to 1P control ($p = 0.046$), and decrease of negative feelings from 1P to 3P control ($p = 0.42$); however, the largely unequal sample sizes make these tests untrustworthy.

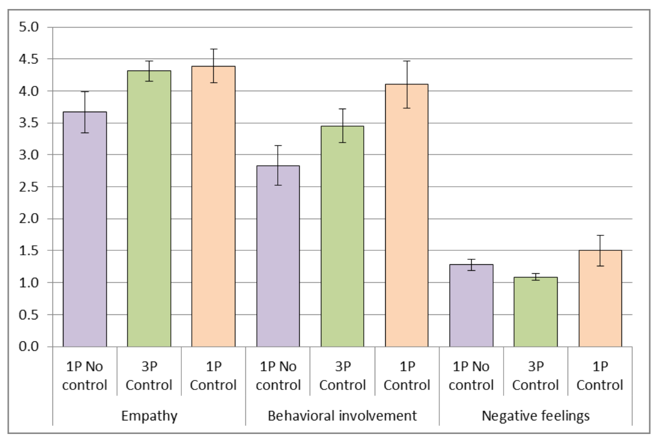

Figure 5. ANOVA results of SP measures by level of control.

Effects of Interaction Level on Game Data

None of the logged game data seemed to have any effect on learning between conditions. However, actions performed on average were more than double in the 1P condition ($M = 66.17$, $SD = 3.43$) compared to 3P ($M = 28.07$, $SD = 12.72$), which was expected since tasks in 3P are distributed between players (corrected t-statistic due to inequality of variances $t(1,17.9) = 10.67$, $p < 0.001$). Trying to do an action at the wrong time was considered an error, except when collaborative action was required, since players were often clicking repeatedly to achieve the right synchronization. Errors made were not found to be significantly different

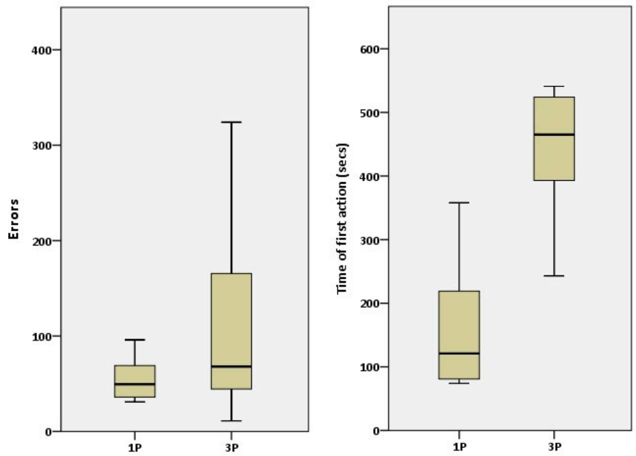

Figure 6. (a) Errors and (b) time of first action per condition.

among conditions, but the distribution of errors was much larger in the 3P condition (1P: min = 31, max = 96; 3P: min = 11, max = 324), as can be seen in a. This can be explained by the different levels of involvement players assumed either due to their game-playing experience and style or fear of making mistakes.

We also noticed from our observations that participants in the 3P condition had trouble figuring out how to troubleshoot the first problem (i.e., either starting up the boiler or attaching a belt to one of the machines). Indeed, students working together took significantly more to carry out the first task compared to the 1P condition ($t(1,31) = -7.18$, $p < 0.001$), as can be seen in b. This is probably an indication of the overhead that collaboration entails, demanding more effort to establish effective communication. However, this extra time for the 3P group did not seem to affect the overall completion time (1P: M = 36mins 52secs; 3P: M = 38mins 57secs).

Path Analysis

In an attempt to explore the combined effect of the independent variable (interaction level) and the self-reported measures (game experience, presence, social presence) to predict the outcome (learning), we decided to use the partial least squares (PLS) statistical analysis method. PLS is comprised of a structural model with the relationships between the latent constructs (blue circles) and a structural equation model with the predictive relationships between each latent construct and its associated observed indicators (yellow boxes) [12]. We used SmartPLS to setup and test the path model we constructed from the literature and our hypotheses. shows the model and the results of the analysis using the same number of cases (N=47) and bootstrapping with 5K samples. Using this model, interaction level significantly predicted learning in the presence of the other constructs ($t = 2.24$, $p < 0.05$), with higher levels of interactivity predicted to produce greater learning. Interaction level also significantly predicted social presence ($t = 2.06$, $p < 0.05$), but not game experience or presence.

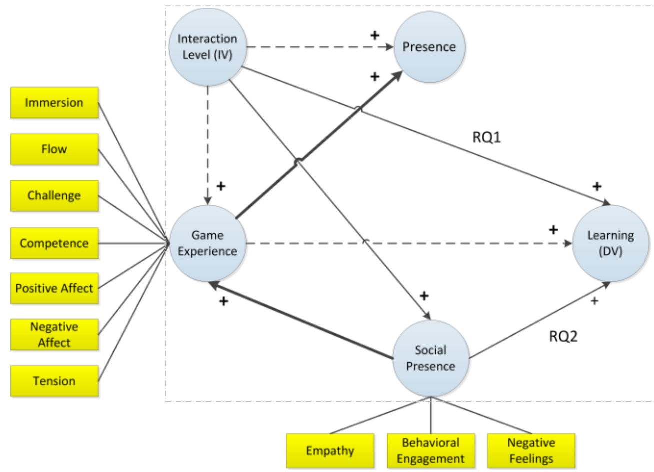

Figure 7. Path model with hypotheses (+ sign indicates positive correlation) and results (lines). Solid lines are 5% significant; thick lines are 1% significant; dotted lines are not significant.

Interestingly enough, this model shows that social presence had a highly significant effect not only on game experience but also on learning. These results support our hypotheses about the positive impact of co-located gameplay and collaboration on game experience and eventually learning.

DISCUSSION

Although the main analysis was not able to reject the null hypothesis, the path analysis did support the combined effect of all the constructs (interaction level, game experience, and social presence) on learning. This result is also supported by research on the impact of increased social presence on enjoyment [16], which has been shown to positively affect learning [23]. This cascade effect of the attributes of collaborative serious games and VEs on learning has been the main premise of our research.

Furthermore, there were many interesting findings that point towards the benefits of co-located collaboration on the overall game experience. Besides the statistical findings, our qualitative results from observations, recordings, and interviews indicate that students enjoyed the 3P condition much more than the other two. There were many cases in the 1P condition where both controllers and non-controllers expressed their preference to engage in a multi-player game. Despite the fact that this was not captured in all our individual GE measures, comparing GE as the mean of all seven measures did reveal a significant difference between conditions.

Increased fun can also be attributed to the simultaneous (or orchestrated) nature of collaboration we have employed in C-OLiVE. The results of comparing controllers in both conditions with 1P non-controllers (Figure 5) indicated a higher interconnectedness felt by the students with control (i.e., behavioral involvement), but also more negative feelings on behalf of the ones with 1P control when compared with 3P players. Although the comparison between 1P controllers and non-controllers did not prove to be significant, it is still interesting to note that 1P controllers felt worse.

This can be justified by the pressure that these students felt by having the fate of the game in their hands, especially if they were not experienced players, as observed during game play but also relayed through the interviews.

In alignment with our finding about the increased time to complete the first task, we have noticed students struggling to establish effective communication during the beginning of the 3P game. Counting also the interventions we had to do in this condition (our policy was to provide a hint from the video if they were stuck for three minutes), the 3P condition demanded more than twice the number of interventions on average compared to 1P (4.6 vs. 2.0). Although this did not seem to negatively affect the game experience of students or the learning achieved in this condition, it does indicate that multiuser social interactions demand greater effort in communicating intent, until commitment demands are revealed to and understood by participants [6].

Despite our effort to conduct as thorough a study as possible, there are some limitations of our experimental design. First, the sample size was small and may not be easily generalizable to a wider population. Second, comparing controllers and non-controllers statistically did not have adequate power, as the sample sizes were largely unequal. Third, the experimental setting was not ecologically valid for the purpose of this research, since it was mainly conducted within controlled conditions. Fourth, the interaction design of C-OLiVE was noticed to be different than commercial games, as commented by some students. Finally, the scientific domain chosen, the collaborative interactions it affords, and the design of these interactions impose some constraints to the overall instructional design; thus we should be cautious in generalizing these findings.

CONCLUSIONS AND FUTURE WORK

We conducted an experimental study to assess the learning benefits that derive from co-located collaboration using a gaming simulation of olive oil production. We ran the study with middle school students in summer camps and schools and found that the increased social presence and enhanced game experience afforded by the simultaneous control of the game eventually facilitate learning. Our observation and interview data verify that students playing together in the 3P condition were more engaged with the game, which created more positive affect but also a sense of interconnectedness. Our analysis also revealed that students playing in this condition had a significantly better understanding of the collaborative tasks compared to their 1P counterparts.

Although we understand that this was a preliminary study and more data are needed to support our hypothesis, we still believe that the findings point towards some interesting direction for CVEs design. More specifically, a multiplayer design should be considered when the time and space are appropriate, as the social interactions, improved game experience, and increased engagement afforded by co-located collaboration can be beneficial for learning.

Although learning is important in museums, we should not underrate their entertainment value for the visitor experience. Thus, our next step includes running the experiment in a museum, in order to investigate how the motivations of a museum visit but also the prior knowledge and interests of the visitors might affect learning [13]. Thus, we will run the study in a museum in Greece, a place with rich cultural history in olive oil production, and consequently anticipated greater prior knowledge on behalf of the students. Moreover, this will allow us to test our hypotheses in an ecologically valid setting, with a much larger sample size that will increase the power and generalizability of the results.

ACKNOWLEDGMENTS

We acknowledge the support of Dane Webster and Reza Tasooji in modelling part of the virtual world. We thank Eric Ragan, Kumbirai Madondo, and Sandra Pöschl for their valuable comments on the experimental design. We also thank the Cultural Foundation of Piraeus Bank in Greece for their consulting and content provision support. Finally, we thank all the participants and their parents.

REFERENCES

1. Allen, S. Designs for learning: Studying science museum exhibits that do more than entertain. *Science Education 88*, S1 (2004), S17–S33.
2. Amory, A., Naicker, K., Vincent, J., and Adams, C. The use of computer games as an educational tool: identification of appropriate game types and game elements. *British Journal of Educational Technology 30*, 4 (1999), 311–321.
3. Barab, S., Pettyjohn, P., Gresalfi, M., Volk, C., and Solomou, M. Game-based curriculum and transformational play: Designing to meaningfully positioning person, content, and context. *Computers & Education 58*, 1 (2012), 518–533.
4. Bellotti, F., Berta, R., De Gloria, A., D'ursi, A., and Fiore, V. A serious game model for cultural heritage. *Journal on Computing and Cultural Heritage 5*, 4 (2012), 1–27.
5. Biocca, F., Harms, C., and Burgoon, J.K. Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence: Teleoperators & Virtual Environments 12*, 5 (2003), 456–480.
6. Brown, B. and Bell, M. CSCW at play: 'there'as a collaborative virtual environment. *Proceedings of the ACM conference on Computer-Supported Collaborative Work (CSCW'04)*, (2004), 350–359.
7. Crook, C., Harrison, C., Farrington-Flint, L., Tomás, C., and Underwood, J. *The impact of technology: value-added classroom practice: final report*. British Educational Communications and Technology Agency (BECTA), 2010.
8. Csikszentmihalyi, M. *Flow: The Psychology of Optimal Experience*. Harper Perennial, London, UK, 1990.

9. Dede, C., Salzman, M., Loftin, R.B., and Ash, K. Using virtual reality technology to convey abstract scientific concepts. In M.J. Jacobson and R.B. Kozma, eds., *Learning the Sciences of the 21st Century: Research, Design, and Implementing Advanced Technology Learning Environments.* Lawrence Erlbaum, Hillsdale, NJ, 2000.
10. Dede, C. The evolution of constructivist learning environments: Immersion in distributed, virtual worlds. *Educational technology 35*, 5 (1995), 46–52.
11. Dede, C. Immersive interfaces for engagement and learning. *Science 323*, 5910 (2009), 66–9.
12. Fair, J.F., Ringle, C.M., and Sarstedt, M. PLS-SEM: Indeed A Silver Bullet. *The Journal of Marketing Theory and Practice 19*, 2 (2011), 139–152.
13. Falk, J.H. and Dierking, L.D. *Learning from museums: visitor experiences and the making of meaning*. AltaMira Press, Walnut Creek, CA, 2000.
14. De Freitas, S. and Griffiths, M. The Convergence of Gaming Practices with Other Media Forms: What Potential for Learning? A Review of the Literature. *Learning, Media and Technology 33*, 1 (2008), 11–20.
15. De Freitas, S.I. Using games and simulations for supporting learning. *Learning, Media and Technology 31*, 4 (2006), 343–358.
16. Gajadhar, B.J., Kort, Y.A.W. De, and Ijsselsteijn, W.A. Shared Fun Is Doubled Fun: Player Enjoyment as a Function of Social Setting. In P. Markopoulos, ed., *Fun and Games, LNCS 5294*. 2008, 106–117.
17. Ijsselsteijn, W., Kort, Y. De, Bellotti, F., and Jurgelionis, A. Characterising and Measuring User Experiences in Digital Games. *Proceedings of the international conference on Advances in Computer Entertainment Technology (ACE)*, (2007), 27–30.
18. Jackson, R.L., Taylor, W., and Winn, W. Peer collaboration and virtual environments: a preliminary investigation of multi-participant virtual reality applied in science education. *Proceedings of the 1999 ACM symposium on Applied computing*, (1999), 121–125.
19. Kirriemuir, J. and McFarlane, A. *Literature Review in Games and Learning Literature Review in Games and Learning*. Bristol, UK, 2004.
20. Klemmer, S.R., Hartmann, B., and Takayama, L. How Bodies Matter: Five Themes for Interaction Design. *Proceedings of ACM Designing Interactive Systems Conference (DIS'06)*, (2006), 140–149.
21. De Kort, Y.A.W. and Ijsselsteijn, W.A. Digital Games as Social Presence Technology: Development of the Social Presence in Gaming Questionnaire (SPGQ). *Proceedings of PRESENCE*, (2007), 195–203.
22. Lave, J. and Wenger, E. *Situated Learning. Legitimate peripheral participation*. University of Cambridge Press, Cambridge, UK, 1991.
23. Malone, T.W. and Lepper, M.R. Making learning fun: A taxonomic model of intrinsic motivations for learning. *Aptitude learning and instruction III Conative and affective process analysis*, 1987, 223–253.
24. Prensky, M. *Digital Game-Based Learning*. McGraw-Hill, New York, USA, 2001.
25. Prensky, M. Digital Natives, Digital Immigrants. *On the Horizon 9*, 5 (2001), 1–6.
26. Resnick, M. Rethinking learning in the digital age. In G. Kirkman, ed., *The Global Information Technology Report: Readiness for the Networked World.* Oxford University Press, Oxford, UK, 2002, 32–37.
27. Rieber, L.P. Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational technology research and development 44*, 2 (1996), 43–58.
28. Roussos, M., Johnson, A., Moher, T., Leigh, J., Vasilakis, C., and Barnes, C. Learning and Building Together in an Immersive Virtual World. *Presence: Teleoperators and Virtual Environments 8*, 3 (1999), 247–263.
29. Sallnas, E. Effects of Communication Mode on Social Presence, Virtual Presence, and Performance in Collaborative Virtual Environments. *Presence: Teleoperators & Virtual Environments 14*, 4 (2005), 434–449.
30. Slater, M. and Schroeder, R. Small-Group Behavior in a Virtual and Real Environment. *Presence: Teleoperators and Virtual Environments 9*, 1 (2000), 37–51.
31. Snowdon, D., Churchill, E.F., and Munro, A.J. Collaborative Virtual Environments: Digital Spaces and Places for CSCW. In D. Snowdon, E.F. Churchill and A.J. Munro, eds., *Collaborative Virtual Environments*. Springer Verlag, 2000, 1–34.
32. Squire, K. and Jenkins, H. Harnessing the power of games in education. *Insight 3*, 1 (2003), 5–33.
33. Squire, K. Changing the game: What happens when video games enter the classroom. *Innovate: Journal of online education 1*, 6 (2005), 1829–1841.
34. Steinkuehler, C. and Alagoz, E. Out-of-School Virtual Worlds Based Programs: A Cross-Case Analysis Methods. *Proceedings of the 9th International Conference of the Learning Sciences (ICLS)*, (2010), 304–305.
35. Winn, W., Windschitl, M., Fruland, R., and Lee, Y. When does immersion in a virtual environment help students construct understanding. *Proceedings of the International Conference of the Learning Sciences, ICLS*, (2002), 497–503.
36. Winn, W. Current Trends in Educational Technology Research : The Study of Learning Environments. *Educational Psychology Review 14*, 3 (2002), 331–351.