Do Low Cost Virtual Reality Devices Support Learning Acquisition?

A comparative study of two different VR devices

Paloma Díaz, Telmo Zarraonandía, Mónica Sánchez-Francisco, Ignacio Aedo, Teresa Onorati[†]

Interactive Systems Group (DEILAB). Computer Science Department

Universidad Carlos III de Madrid

Leganés, Madrid, Spain

{pdp, telmoz, mosanche, tonorati}.inf.uc3m.es, aedo@ia.uc3m.es

ABSTRACT

Virtual Reality (VR) can support engaging and effective learning and current VR devices commercially available are opening up opportunities and expectations in higher education. An open research question is determining whether the fidelity of the VR device (e.g. resolution, framerate, etc.) impacts learning. This work aims to address this question by conducting a study wherein the quality of the experience and the knowledge acquired using a high fidelity (and high priced) VR headset is compared against a lower fidelity (and lower priced) VR headset. The results of the study do not reveal significant difference in spatial and experiential learning. The user satisfaction and experience are also similar regardless of the headset used.

CCS CONCEPTS

• Human-centered computing → Virtual reality; User studies;

KEYWORDS

Virtual reality, wearables devices, empirical study, learning affordances.

1 Introduction

Virtual Reality (VR hereafter) aims to augment human capabilities in different kinds of activities including learning [22]

[2]. The use of head mounted displays (HMD) in VR learning environments adds a higher sense of immersion and presence to support learning affordances such as situated learning, and increase motivation and engagement [5, 7]. In the last years, HDMs have become commercially available as low cost wereables that make use of the user mobile phone like Google Cardboard [11] or Samsung VR Gear [22], or as higher fidelity and higher cost VR systems such as HTC Vive [28] and Oculus Rift [20]. This variety of devices feeds the expectations about the future of an immersive VR not confined anymore to labs but that will be experienced almost everywhere. However, according to the Gartner's Hype Cycle of Emerging Technologies, VR hasn't reached yet the "plateau of productivity" [9]. In other words, it has not reached a mass adoption status. One of the reasons might be precisely that such variety of options makes it difficult to choose the right one to invest in. From an educational point of view, there is no doubt that one of the challenges for VR being adopted has always been the cost [6]. VR equipment prices vary from less than $100 to over a thousand dollars. In addition, new technologies have to suit not only institutions' budgets but also their learning goals. Thus, new technologies need to be affordable and provide educational value in order for education administrators to make investments.

In this context, this research aims at exploring whether low cost VR devices provide comparable learning outcomes to higher quality/higher cost VR devices. With this purpose we conducted an experiment comparing the learning outcomes of two specific VR commercial wearables with very different prices and technical capabilities: the Oculus Rift VR headset that cost around $500 and provides a vivid experience and an increased sense of immersion; and the Google Cardboard that makes use of the user own phone to support low quality 3D immersion but costs less than 8$. Building upon previous studies that question the influence of immersion in learning acquisition compared to the contribution of embodied interactions [19][21], our hypothesis is that the cheapest gear can still be valid when a high level of realism is not required to attain learning goals. To test this hypothesis, we run an empirical study with 41 participants, where the devices were used to analyze learning acquisition in two learning tasks: spatial and experiential learning. Spatial learning implies acquiring conceptual knowledge by interacting with a representation of a

real or abstract concept in a virtual world [26][19]. Experiential learning is the process of acquiring procedural knowledge by doing or experiencing a process [5][16]. These two learning tasks were selected since they are the two types of learning activities which are improved by using VR as discussed in [15]. Moreover, since VR devices are starting to enter our living spaces, specially for gaming, we assessed how users perceived the experience, to understand whether the lower quality device might be deceiving or not. According to our results, they were no significant differences in learning gain nor in the perception of the user experience.

The remaining of the paper starts by reviewing related works to frame the problem. Then the VR experience used for the experiments is described focusing on the type of learning tasks proposed. The empirical experiment is described in section 4, followed by a discussion on the main findings. Conclusions and future areas of research are drawn at the end of the paper.

2 Related Works

VR or 3D worlds have been used in education since the 90s, whether as desktop interfaces or as immersive simulations using HMDs or CAVE systems [2][21][17]. This type of technology is particularly adequate to support learning not only because it is intrinsically motivating and engaging, but also because it has a number of learning affordances [7][19][5]. Thus, Dede identifies three potential affordances of well- designed immersive VR [7] they improve learning through the exploration of knowledge from multiple perspectives, and they enhance both situated learning and the transfer of knowledge to the real world thanks to the possibility of recreating real scenarios that learners can interact with. Learning affordances rely upon three key features of VR: immersion, fidelity and a higher level of learners' active participation [13]. From these three features, immersion and fidelity strongly depend on the technical features of the VR device. As mentioned earlier, there are a variety of VR devices that make it possible to interact with VR worlds. Each of them has a different level of fidelity and immersion. However, several studies [7][21][4] conclude that a high level of immersion might not be decisive to reach the learning goals. Other features of VR, like the physical interaction with the virtual space or the capacity to explore concepts from multiples perspectives [6][7] can have a greater influence in the learning outcomes. Under these premises, this paper explores how two different commercial VR devices perform in a learning situation where immersion is not a key factor. In particular, our hypothesis is that low cost VR can be used to support spatial and experiential learning offering useful experiences to learners.

3 Assessing the learning affordances of two different VR devices

In order to test whether VR devices with low levels of realism and immersion do still support learning acquisition we conducted an experiment that compares the learning experience of two groups of participants using 2 very different devices in terms of immersion quality and cost: the Oculus Rift and the Google Cardboard. The study is aimed at assessing whether the economically affordable option is still valid to acquire knowledge in certain situations where fidelity and immersion can be sacrificed to reach the learning goals [21] and how the experience is perceived by learners.

The virtual environment chosen is an automobile repair shop with a car in the center of the room waiting to be explored and fixed. The user can walk around using the gamepad joystick, examine the car closer and interact with the different objects in the room, including the car components. This scenario was chosen since it does not require a high-level of realism and immersion as a training system for neurosurgeons does. However, it is still an example of situated learning [7][26] since a real world example is used to acquire spatial and procedural knowledge and skills. The main goal of the experiment was testing whether both devices could be used to implement two types of learning tasks that according to [5] are two basic learning affordances of VR: spatial and experiential learning. Indeed, the car engine can be explored to learn about its components, their properties or their position (conceptual and spatial knowledge gained through spatial learning) or how they can be manipulated to fix the engine (procedural knowledge gained through experiential learning). Next subsections describe the two learning scenarios of the study.

3.1 First learning scenario: spatial learning task

Spatial learning is a key activity in which learners explore and interact with a representation in the virtual world (real or abstract object or process) to acquire some conceptual knowledge [19]. The objective of the first learning task in the study was to help the user to gain an understanding of the main engine components, so that she could carry out some basic car maintenance operations, such as checking the oil level, fill the brake fluid tank, clean battery contacts, etc. More specifically it is aimed at supporting two learning objectives:

- To identify and recognize the main components of a car engine, that is to acquire spatial knowledge on the car engine's structure.

- To understand their principal characteristics and functionalities, that is to acquire some conceptual knowledge on the car engine's functioning.

At the beginning of the activity the whole chassis of the car disappears to allow the user to have a clear view of its internal mechanism. A panel displays names of different components of the engine and suspension systems, and the user has to identify them in the car by exploring the engine (see Figure 1). The user can select the components of the car using a common interaction solution in VR environments: raycasting the user's gaze. A crosshair marks the center of the user's vision. When the user looks and places the crosshair on top of the component whose name is displayed in the panel, the object is highlighted by changing its color. The user then presses the X button on the gamepad, and a green panel confirms that the answer is correct

and displays information about the component's main characteristics and functionalities. If the user is unable to find the car component requested, he/she can ask for help by pressing tthe corresponding button on the gamepad, and the component will be automatically highlighted. This process is repeated for a total of 13 component parts, which include the radiator, the generator, the thermostat, the brake fluid tank, the oil cap and the positive and negative terminals of the battery, among others.
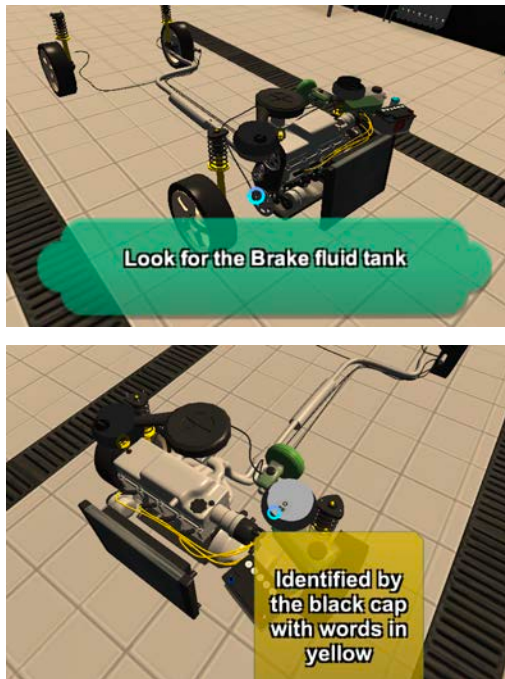

Figure 1. Identifying parts of the engine in the VR Environment during the spatial learning task

3.2 Second learning scenario: experiential learning task

Experiential learning or action learning is the process of learning by doing or experiencing [14][16]. VR makes it possible to simulate environments where experiential learning is cheaper, safer and easier to personalize than real world experiences [6][7][5][21]. In our case, the second learning task was designed as an experiential learning task aimed to help the learners acquire some procedural knowledge on the car mechanics. The task proposed was to change a car battery. This is a key procedure for the car maintenance that has to be carried out in a very specific order to prevent damaging the car engine.

At the beginning of the task a panel displayed the lists of steps to follow, which involved the disconnection and removal of some car components, and their re-connection and installation of new ones. Once the learner indicated she had memorized them, the panel disappeared and she was asked to carry out the replacement operation. To help simulating the operations required for the task a button-bar was displayed on the top of her view during the task. Initially, the bar displayed one single button representing the new

battery to install. To disconnect or remove a car component, the user needed to highlight it by setting the gaze-crosshair on top of it and clicking the X button on the gamepad. The component would then disappear from the car, and its name would be added to the button-bar (see Fig. 2). To re-connect a component or install a new one the user just had to select its name from the button-bar, which could be navigated using the gamepad buttons. If the user failed to execute the steps in the right order, missed one of them, or made any other kind of mistake, an animation displaying smoke coming from the engine was activated to indicate the error, and the exercise had to be repeated from the beginning.
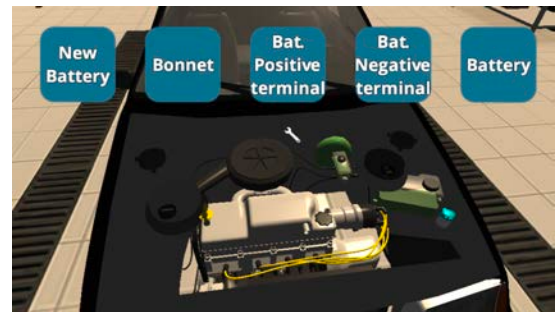

Figure 2. Learning to replace the battery in the exploratory learning task

4 Experiment

We conducted a user study to assess whether the cheaper and low-quality of immersion device can still support learning acquisition in both learning scenarios. To explore this research question, we run the experiment with two Virtual Reality (VR) devices, with a different grade of immersion and presence [30][17]: Oculus Rift and Cardboard Glasses. In this paper we consider Oculus Rift as a higher quality immersion device due to the fact that relevant parameters to support presence like the refresh rate or the field of view are much better than those of the Cardboard.

The experiment involved a total of 41 participants and served to investigate two main hypotheses derived from previous studies:

- H1: The low-fidelity/low cost device supports spatial learning in a similar way than the high-fidelity/high cost device.

- H2: The low-fidelity/low cost device supports experiential learning in a similar way than the high-fidelity/high cost device.

Also, we explored a third hypothesis to understand how the low-cost experience was perceived by the participants, since the acceptance of a technology does not only depend on objective factors like cost, quality or performance but also on subjective perceptions and beliefs about the technology [1][27].

- H3: The low-fidelity/low cost device provides learners with a worst experience than the high-fidelity/high cost device

Since getting used to interacting with a VR HMD is not easy and requires some training, we applied a between-subject design in which each participant was exposed to only one condition [18], that is, each of them used only one device for both tasks. This kind of approach makes studies shorter and avoids fatigue or frustration [18]. Also, considering we wanted to measure the perception of the experience, being exposed to higher quality might have biased the opinion of participants to choose always the best quality without thinking if lower quality would be enough. The following subsections describe the experiment, starting by the participants and technical apparatus used, the procedure followed, the data collection mechanisms and the main results obtained.

4.1 Participants and Apparatus

41 people (14 female and 27 male) aged 17 to 26 (M = 19.8, SD= 2.22) took part in this study. The participants of the experiment were distributed in two groups: G1 used the Oculus Rift, and G2 used the Cardboard Glasses. G1 included 20 subjects (6 female and 14 male) aged 18 to 24 (M = 21.3, SD = 1.58) and G2 21 subjects (8 female and 13 male) aged 17 to 26 (M = 18.52, SD = 1.86). We recruited the participants from the first year courses of two grades of the Universidad Carlos III de Madrid: computer science (G1) and biomedical engineering (G2). Due their age, we expected the proposed activity would be of special interest, as most people of their age have either taken the driven license exam recently or they might be thinking of doing it. When asked to rate their command on car mechanics using Likert scale ranging from 1 to 7, results showed that most recognized to be somehow familiar but not experts (M = 2.70, SD =1.19).

G1 performed the experiment with an Oculus Rift in an Apple i-Mac computer. G2 run the experiment with a Google cardboard kit v2 Minikanak, compatible with a 4.8 screen inch Android phone, a Samsung Galaxy S III (see Fig. 2). The participants of the both groups used a PlayStation DualShock 3 wireless controller to navigate and control their actions in the VR environment. In order to have exactly the same type of interaction during the experiment as G1, the Cardboard kit used in G2 incorporated a head strap and nose pad, with the purpose of leaving participants' hands free to use the same gamepad in both groups.

Figure 2. Participant during the experiment

4.2 Data Collection Mechanisms

As stated earlier, the experiment is a between-subjects design. The independent variable (quality of the immersion) is due to the different features of the two HMDs used. The dependent variable (learning outcomes) is analyzed using quantitative and qualitative data. Quantitative data are related with the learning outcomes of both tasks. To assess the learning outcomes in the spatial learning scenario we used pre-test and a post-test knowledge tests as in [3], that asked participants the same ten questions before and after exploring the virtual world. 6 of the questions aimed to assess the capacity of the users to identify and recognize components usually involved in car maintenance operations, while the other 4 evaluated their knowledge about the function and purpose of some of the main parts of the engine. For each subject we calculate a Total Score as follows: being PoTn and PrTn the scores obtained for question n, the learning gain for question n is defined as $Ln = (PoTn - PrTn)$.

As for the second task, the experiential learning scenario, since it involves procedural knowledge and procedural knowledge is difficult to put into words, no pretest was used. Instead of, participants, all of whom admitted not to know how to replace a battery before the experiment, were asked to describe the sequence of steps after interacting with the simulation using a retention test in a similar way as in [10].

To assess the user perception on the quality of experience (H3) we build upon the conception of experience by Dillon who differentiates between the *process*, the *outcomes* and the *affect* [8]: *process* assessment is related with measuring aspects of the interaction process between the user and the device; the *outcomes* are related to what the user gains with the interaction process, and, finally, the *affect* is concerned with understanding the emotions and attitudes promoted by the interaction process.

In our case, the *process* was measured using two parameters: the time required to perform the learning tasks and the number of user's errors.

Outcomes were measured with the pre and post tests used to validate H1 and H2 that already collect whether the interaction makes it possible to acquire conceptual and procedural knowledge through the two learning scenarios proposed.

Finally, to understand how participants perceived the experience and their attitude towards the technology (*affect*) we used an Experience Test filled at the end of the session. Most questions of the test are based on the Presence Questionnaire introduced in [21]. We considered this test suitable for the study since it focuses on subjective impressions about the experience as discussed in [25]. We excluded the items from the questionnaire regarding features not implemented in this specific experience, as sounds or haptic stimulation, and selected 10 questions (from Q1to Q10 in Table I) that explored three aspects of the experience usually regarded to enhance learning and performance: the level of involvement [10], the distraction [24] and the level of control [29] [13]. A last question explored how much participants liked the experience (Q11 in the Table). All these questions were measured using a scale ranging from 1 to 7, where the two end points where semantically tagged to facilitate the interpretation of the grading (see Table I). Additionally, four open questions collected

information on the user perception and attitude towards the technology: two of them asked participants to indicate the aspects they liked the most as well as the limitations they experienced, and the other two explored whether they would use that system and which potential uses of VR they foresee.

4.3 Procedure

The experiment consists of seven different phases. The first phase was a pre-test done to better understand the participants' demographics as well as to assess their previous knowledge on the topic to be explored (car mechanics). Then, a member of the research team (evaluator hereafter) trained the subjects on how to navigate through and interact with the virtual environment using the gamepad controls. The goal of the experiment is also introduced to remind participants they are not being evaluated, not on their knowledge on car mechanics neither on their ability to use the gamepad controls Once participants were ready to interact with the virtual world, they started by executing the spatial learning task. The evaluator stayed with the participants just in case they asked for assistance, but she did not interfere with the interaction process. Once the learning task was finished, the participant filled a post knowledge test that was used to assess the learning gain. As in similar experiences, learning gain was measured by comparing the pre and post test scores [12][3]. Next, they carried out the experiential learning task, and they filled in a new post-test to measure the knowledge gained during this second task. Finally, the participants completed the questionnaire about the overall experience (see Fig. 3).
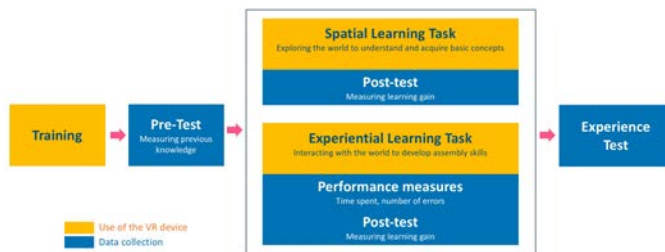

Fig. 3. Experiment procedure

4.4 Results and Findings

In this section we discuss the results and main findings focusing first on the learning outcomes (hypotheses H1 and H2), and then on the perception of the experience (hypothesis H3).

Learning outcomes. With regards to the spatial knowledge task, all participants scored better in the post-test. A Paired samples t-test was conducted to evaluate whether there was significant learning in the two groups. The score for G1 in the pre-test was M = 5.15, SD = 1.55, and post-test M = 7.7, SD = 1.61, with t(20) = -5.59, p < 0.0001. The score for G2 in the pre-test was M = 5.23, SD = 2.38 and the post-test M = 7.8, SD = 1.67, with t(21) = -5.51, p < 0.0001.

Figure 4 depicts the boxplots with the means and distributions of the scores obtained in the pre and post-tests. As shown in the

picture the final results in both groups were very similar after the educational intervention. However, G1 performed slightly worse in the pre-test. In order to measure the differences in Learning Gain (L) after the activities and independent-samples t-test was conducted. In this case there was not a significant difference in the scores of L for G1 (M = 3.45, SD = 1.82) and G2 (M = 2.57, SD = 2.08) conditions, t(41) = 1.39, p = 0.17. Figure 5 depicts graphically the differences in L between the two groups.

These results suggest that there was significant learning in both groups for the spatial learning task and there was no significant difference in the learning gain between the two groups, so that hypothesis H1 can be considered as validated.

Concerning the experiential learning task, before starting the activity most participants of both groups reported not knowing how to change the battery (17 subjects in G1 and 20 in G2), or just having only a slight idea of the process. Regardless the device used all the participants behaved similarly, and no significant delays or problems to complete the assemble of the battery were detected in any of the two groups. When asked to describe the steps to follow at the end of the experiment, only 1 participant in G1 provided an incomplete answer and 3 participants in G2 responded incorrectly.
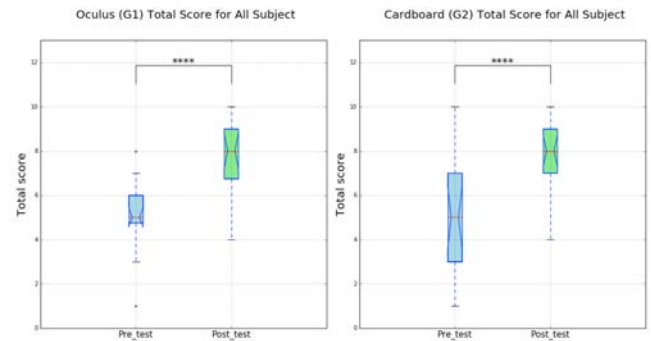

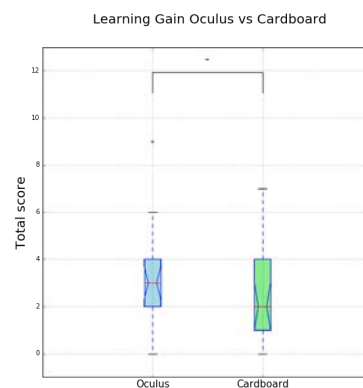**Fig. 4. Test scores in the spatial learning task for G1 (left) and G2 (right)**

Fig. 5. Differences in the learning gains in the spatial learning task for the two groups

User Experience. As explained before, the user experience was measured collecting information on the process and the affect generated by the interaction with the VR worlds. Concerning the process, we measured the time spent and the errors performed. Figure 6 depicts the differences in total time taken to complete the spatial and experimental tasks between the two groups. As shown in the pictures, in both cases the participants of G2 required more time to complete the activities. An independent-samples t-test was conducted to compare the results for the spatial task G1 (M = 3.37, SD = 1.19) and G2 (M = 5.22, SD= 1.10), with t(41) = 5.01, p =<0.0001. This result suggest that the device had influenced the performance time. To test the execution time for the procedural task we follow the same method, and the results for G1 (M = 1.10 min, SD = 1.79), and G2 (M = 2.29 min, SD = 1.28), a t(41) = 1.24 and p = 0.21 show not significance for the difference in time spent for the experiential learning task. Participants didn't have significant errors using the VR worlds in both cases.

To collect information on the emotions and attitudes generated during the experience, what is called affect in [8], we used the user experience questionnaire whose answers are summarized in Table I. As shown in the table, the means of the responses collected are very similar in the two groups, with only small
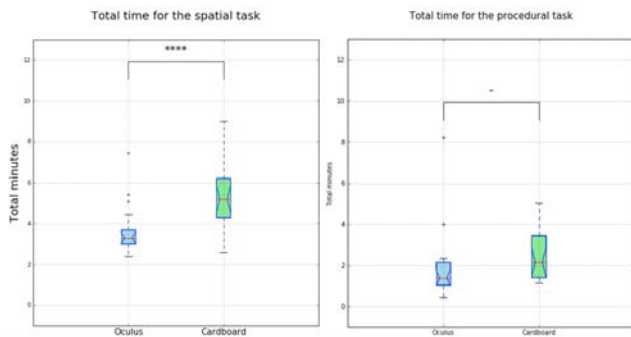

Fig. 6. Total time spent with each device during the execution of spatial task (left) and procedural task (right)

differences of some decimals between the two groups. When obtaining the total score for each of the three different factors of presence considered in the questionnaire (reverse item Q9), the means of the total scores also resulted very similar (Table II). In a range from 3 to 21 the total score for the grade of involvement in G1 was M = 16.25, SD = 2.65 and for G2 M= 16.71, SD = 2.07. In the case of distraction in a range from 4 to 28 the total score for G1 was M = 14.85, σ = 3.59 and for G2 M = 14.81, σ = 3.80. Finally, for the control in a range from 3 to 21 the total score for G1 was M = 16.65, SD = 2.01 and for G2 M = 16.47, SD = 2.50. A Shapiro-Wilk test indicates that the data is not normally distributed for none of the three factors. The Mann-Whitney U test showed that the differences between the G1 and G2 scores for involvement, distraction and control are not statistically significant (Table II).

With regards to the last question (Q11), that the participants used to rated how much they liked the experience, the responses were

TABLE I: PQ QUESTIONS AND RESULTS.

		G1		G2	
		M	SD	M	SD
INVOLVEMENT					
Q1	How much did the visual aspects of the environment involve you? (1 – Not involved / 7 – Very involved)	5.75	0.71	5.57	1.02
Q4	How compelling was your sense of objects moving through space? (1 - Not compelling/ 7 – Very compelling)	5.3	1.45	5.28	1.19
Q10	Were you involved in the experimental task to the extent that you lost track of time? (1 –Not involved / 7 – Very involved)	5.2	1.20	5.95	0.97
DISTRACTION					
Q2	How aware were you of events occurring in the real world around you? (1 - Not aware / 7 – Very aware)	4.0	1.71	4.04	2.08
Q3	How aware were you of your display and control devices? (1 - Not aware / 7 – Very aware)	4.9	1.37	4.61	1.65
Q8	How much did the visual display quality interfere or distract you from performing assigned tasks or required activities? (1 – Not interfere / 7 – Very interfere)	3.25	1.20	3.47	1.32
Q9	How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities? (1 –Not concentrate/ 7 – Very concentrate)	5.3	0.80	5.33	0.96
CONTROL					
Q5	How much did your experiences in the virtual environment seem consistent with your real-world experiences? (1- Not consistent / 7 – Very consistent)	5.1	0.90	5.28	1.55
Q6	How quickly did you adjust to the virtual environment experience? (1- Not quickly / 7 – Very quickly)	5.7	1.49	5.57	1.12
Q7	How proficient in moving and interacting with the virtual environment did you feel at the end of the experience? (1 - Not proficient / 7 – Very Proficient)	5.85	0.98	5.61	1.07
SATISFACTION					
Q11	Please, rate how much did you like the experience (1- Very Little / 7 – Very much)	6.40	0.68	6.47	0.67

TABLE II: SUMMARY OF RESPONSES FOR THE THREE FACTORS RELATED WITH SUBJECTIVE IMPRESSION OF PRESENCE

	G1		G2		
	M	SD	M	SD	MANN-WHITNEY U
INVOLVEMENT	16.25	2.65	16.71	2.07	*U=198.5, p = .762*
DISTRACTION	14.85	3.59	14.81	3.80	*U=203.5, p = .865*
CONTROL	16.65	2.01	16.47	2.50	*U=212.5, p = .947*

also similar and very high for the two groups, M = 6.40, SD = 0.68 for G2 and M = 6.47 , SD = 0.67 for G2.

Finally, the responses of the participants to the open questions do not allow to point out great differences in terms of experience and user satisfaction between the two groups. When asked about the aspects of the activity they liked the most, many participants, regardless of the device used, highlighted the realism of the simulation, (7 participants in G1 and 5 participants in G2) with comments like *"… I felt I was there"* or *"…To have the sensation that I was really moving around the virtual space"*. Also, participants frequently mentioned as a highlight of the experience just the opportunity to use a VR application (6 participants in G1 and 5 participants in G2). Finally, a few subjects (3 participants in G1 and 2 participants in G2) acknowledged as a major outcome some benefits that the VR technology could report for education.

With regards to the aspects of the experience they liked less, the answers were also very similar in both groups. The most repeated negative comments made reference to the quality of the images (5 participants in G1 and 5 participants in G2), and to the application's controls (5 participants in G1 and 5 participants in G2). More specifically, for some subjects it was a bit difficult to use the gamepad without being able to see it, while for others the controls were either too sensitivity or too fast. In addition, 2 subjects in G1 and 3 in G2 mentioned that they felt a bit dizzy at the end of the activity.

When asked whether they would use it in a specific context where this type of knowledge is required, the drivers' license test, all participants agreed that it would be really useful and even cheap. Finally, when asked about potential uses of VR most participants mentioned educational uses (16 in G2 and 15 in G1) and entertainment (6 in G1 and 5 in G2).

5   Discussion

In this section we discuss our main findings concerning the two initial goals: understanding whether learning acquisition and the subjective perception of the experience can be affected by the device quality.

Low-fidelity does not necessarily have an impact in spatial nor experiential learning

Even though the experiment might have limitations in terms of the time spent with the devices or the profile of the participants, the data collected suggest that the two learning hypothesis, H1 and H2, are satisfied since for both types of learning tasks, spatial and experiential, there was a similar learning gain in the two conditions, so we can conclude that:

- The low-fidelity/low cost device can support spatial learning in a similar way than the high-fidelity/high cost device.

- The low-fidelity/low cost device can support experiential learning in a similar way than the high-fidelity/high cost device.

Our results confirm the results obtained in previous studies [19][21] where immersive and non-immersive experiences were compared. In our case, we focused on immersive experiences using quite different levels of quality to assess whether such difference could affect learning in order to provide educators with some criteria to decide which kind of technology they should use. The topic selected in the experiment made it possible to check how participants acquired conceptual knowledge about the car engine by navigating through a 3D representation (spatial learning task) as well as procedural knowledge about the engine functioning by interacting with it to fix the battery (experiential learning task). According to our data, there is no evidence that higher quality impacts the learning gain in any of the two learning tasks. As other studies have concluded, the quality of the graphics or the level of immersion does not affect learning as much as the capability to interact with the world and the quality of the learning tasks proposed. All of our participants valued the fact of being able to get immersed into the virtual space, move around it and interact with the different pieces of the engine to understand how it works. In the words of one of the participants: "*What I liked the most was to have the possibility to see the car components in the surrounding space. It made you feel that the activity of mounting and un-mounting the battery was rea*l". Even when envisioning the future uses of technology, participants in the low-fidelity group highlighted the educational value of VR in different domains.

The only significant difference was in the time required to complete the learning tasks, that was slightly higher in the Cardboard condition. However, this difference might not be relevant enough to untie the situation, except for learning situations where response time in a specific task is critical.

Low-fidelity does not necessarily have an impact in the perception of the quality of the experience.

Given that no significant differences in learning outcomes were perceived in the two conditions, we added a new hypothesis in the third experiment, H3, to assess whether the experience was perceived in a better way using the high quality device. VR devices are starting to enter our homes, especially for gaming, offering entertainment experiences where the quality of the graphics and the level of immersion is quite high and, particularly, higher than the worlds used in this experience. With H3 we aimed at understanding whether participants were somehow deceived given their previous experience or, the opposite, astonished by a new technology they had never tried. With this purpose, the Experience test explored subjective perceptions related with typical aspects related with the perception of the sense of presence (including the level of involvement, the distraction, and the degree of control), the user satisfaction, the intention of use the system and the perception of the technology utility. According to the results there was no evidence that the lower quality condition had any impact on the perceptions and attitudes of the participants towards the technology. Even though some of them recognized as a limitation of the experience the quality of the graphics, then they recognized as a relevant affordance experimenting the sense of being there and, indeed, all the questions related with presence were rated very positively and in a similar way in both groups. That might suggest that users might not demand high quality

representations for learning tasks as they do for other experiences like gaming. In fact, several participants in the G2 groups, low level/cost condition, highlighted they valued very much the possibility of interacting with the car instead of using traditional media like books to understand how the car engine works.

Also, when asked about the intention to use the technology, all participants enthusiastically agreed and among the benefits of the VR technology they stated that learning would be more fun ("… *it was a much more enjoyable way to learn about mechanics*"), effective ("*... an effective way of learning*", "*…it helped more than any written explanation*") and useful to improve concentration *("... it made you disconnect from the real world and concentrate more in what you are doing")*.

6 Conclusions

In this paper we analyzed two different devices with very different qualities and prices to analyze whether they might impact the learning process with a view to provide useful hints to educators and educational administrators about the right technology to invest in. According to the data we collected in the two studies we performed, there is no big difference in terms of learning outcomes, not even in the perception of the utility of the experience. It seems that the low-quality device might be as good as a high quality device since fidelity is less important than the picture superiority, the structuring, and the tuning effects. Educators should focus more on the type of learning activities they propose, the richness of the interaction and the level of challenge than on the quality of the VR world.

For that to be possible, more research on enabling end users (that is, teachers or even students) to create or adapt their own educational VR experiences is required. End users as the owners of the problem know better about the students' needs and capacities, the learning goals pursued, and the possibilities and constraints of the educational context where the technology will be deployed. All this knowledge needs to be considered to design a useful and efficient VR educational experiences and, hence, it would be ideal to integrate end users as developers. However, the technical knowledge required to create a VR world is quite high, for which end-user development tools are required. Such end-user development tools should hide the technical complexities and provide easy to use and domain dependent user interfaces that facilitate the creation of new VR experiences or its adaptation to specific educational tools.

ACKNOWLEDGMENTS

REFERENCES

[1] Ritu Agarwal and Elena Karahanna. 2000. Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage, MIS quarterly, pp. 665-694

[2] Suzanne C. Baker, Ryan K. Wentz, and Madison M. Woods. 2009. Using virtual worlds in education: Second Life® as an educational tool, *Teaching of Psychology*, vol. 36, no. 1, pp. 59-64.

[3] Sarah Buchanan, Jared Bott and Joseph J. LaViola Jr. 2015. The Influence of Multi-Touch Interaction on Procedural Training", in *Proc. of the 2015 ACM International Conference on Interactive Tabletops & Surfaces*, pp. 5-14.

[4] Fabio Buttussi and Luca Chittaro. 2017. Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Transactions on Visualization and Computer Graphics* 2626, c: 1–1. https://doi.org/10.1109/TVCG.2017.2653117

[5] Barney Dalgarno and Mark J W Lee. 2010. What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology* 41, 1: 10–32. https://doi.org/10.1111/j.1467-8535.2009.01038.x

[6] Meredith Bricken. 1991. Virtual reality learning environments: potentials and challenges, ACM SIGGRAPH Computer Graphics, vol. 25, no. 3, pp. 178-184.

[7] Chris Dede. 2009. Immersive interfaces for engagement and learning. *Science (New York, N.Y.)* 323, 5910: 66–9. https://doi.org/10.1126/science.1167311

[8] Andrew Dillon. 2002. Beyond usability: process, outcome and affect in human-computer interactions. Canadian Journal of Library and Information Science, vol. 26, no. 4, pp. 57-69

[9] Gartner's 2017 Hype Cycle for Emerging Technologies 2017: http://www.gartner.com/newsroom/id/3412017, Accessed on: Mar. 7, 2019.

[10] Eleanor Jack Gibson. 1969. Principles of perceptual learning and development, Englewood Cliffs, N.J.: Prentice-Hall

[11] Google, https://vr.google.com/cardboard/, Accessed on: Mar. 7, 2019.

[12] C. Shawn Green, Tilo Strobach and Torsten Schubert. 2014. On methodological standards in training and transfer experiments, *Psychological Research*, vol 78, no 6, pp. 756-772

[13] John Hedberg and Shirley Alexander. 1994. Virtual reality in education: defining researchable issues, *Educational Media International*, vol 31, no 4, pp. 214–220

[14] Leslie. Jarmon, T. Traphagan, M. Mayrath, and A. Trivedi. 2009. Virtual world teaching, experiential learning, and assessment: An interdisciplinary communication course in Second Life, Computers & Education, vol 53, no 1, pp. 169-182.

[15] Paul A. Kirschner. 2002. Can we support CSCL? Educational, social and technological affordances for learning, in *Three worlds of CSCL: can we support CSCL?*, P.A. Kirschner (Ed.), Heerlen, The Netherlands: Open University of the Netherlands, pp. 7-47

[16] David A. Kolb. 2014. Experiential learning: Experience as the source of learning and development. FT press

[17] Péter Tamás Kovács, Niall Murray, Gregor Rozinaj, Yevgeniya Sulema, and Renata Rybárová. 2015. Application of immersive technologies for education: State of the art. *Proceedings of 2015 International Conference on Interactive Mobile Communication Technologies and Learning, IMCL 2015*, November: 283–288. https://doi.org/10.1109/IMCTL.2015.7359604

[18] Jonathan Lazar , Jinjuan Heidi Feng, and Harry Hochheiser. 2017 Research Methods in Human Computer Interaction, Morgan Kaufmann.

[19] Robb Lindgren, Daniel L Schwartz, Robb Lindgren, and Daniel L Schwartz. 2009. Spatial Learning and Computer Simulations in Science. 37–41. https://doi.org/10.1080/09500690802595813

[20] Oculus, http://riftinfo.com. Accessed on: Mar. 7, 2019

[21] Nico Rutten, Wouter R. Van Joolingen, and Jan T. Van Der Veen. 2012. The learning effects of computer simulations in science education. *Computers and Education* 58, 1: 136–153. https://doi.org/10.1016/j.compedu.2011.07.017

[22] Joseph Psotka. 1995. Immersive training systems: Virtual reality and education and training, *Instructional science*, vol 23, no 5-6, pp. 405-431, 1995.

[23] Samsung, http://www.samsung.com/us/explore/gear-vr, Accessed on: Mar. 7, 2019

[24] Shannon F. Harp and Richard E. Mayer. 1998. How seductive details do their damage: A theory of cognitive interest in science learning, Journal of educational psychology, vol 90, no 3, p. 414.

[25] Mel Slater. 2004. How colorful was your day? Why questionnaires cannot assess presence in virtual environments. Presence: Teleoperators and Virtual Environments, vol 13, no 4, pp. 484-493.

[26] Jorge Trindade, Carlos Fiolhais and Leandro Almeida. 2002. Science learning in virtual environments: a descriptive study. British Journal of Educational Technology, vol 33, no 4, pp. 471-488

[27] Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis and Fred D. Davis. 2003. User acceptance of information technology: Toward a unified view, MIS quarterly, pp. 425-478

[28] Vive, https://www.vive.com/eu/. Accessed on: Mar. 7, 2019

[29] Denise Whitelock, Paul Brna, and Simon Holland. 1996. What is the value of virtual reality for conceptual learning? Towards a theoretical framework. *Proceedings of EuroAIE*

[30] Bob G Witmer and Michael J Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments,* 7, 3: 225–240. https://doi.org/10.1162/105474698565686