Can Different Instruction Methods and Previous Experience Affect User Performance in a Virtual Reality Environment?

S. ADOLFSSON, E. BENNERHED, C. BÖHM, S. RÖNNBERG & E. STAAF
2017-06-02

Abstract

In recent years, virtual reality has become increasingly important as a tool for learning, however, there are still issues that need further research. For example, how its attributes can support learning or how to achieve more effective learning when using this technology. Therefore, we chose to investigate which method of instruction that is best suited for a virtual environment and how previous experience in gaming or virtual reality affect task performance. A total of 34 participants took part in the study, and there were two condition where the user either followed instructions based on text or visual objects. The task was to stack cubes on top of each other in the same order as the instructions showed. The results were evaluated both qualitatively and quantitatively. The result showed no significant difference between text and visual object, and there was no significant difference between participants with previous experience with gaming or virtual reality compared to participant with non or less experience. However, we did find that participants used different hand techniques, where gamers used more two handed strategies. Yet, one handed strategies resulted in higher precision. We also found that repetition and the amount of practice are important variables.

Introduction

Today it is possible to explore environments by using 360-degree imagery. Environments like this, such as virtual environments, are created for human's curiosity to explore beyond reality, and to navigate and interact with (Bates-Brkljac, 2012). Although virtual reality is recognized as a tool for learning, there is still issues that need further research (Chen, 2006).

A study made by Healy et al. (2013) explored which modality that is best suited for presenting instructions, but only using a computer display. They compared reading messages, hearing instructions, seeing how to execute a task and a combination of them. The results showed that seeing had the best execution accuracy and reading the worst. However, Healy et al. (2013) found that regardless of which type of instruction that was used, there were no consistently large effect on receiving and following instructions. Hence, they claim that the most important variables are repetition and the amount of practice.

Furthermore, Warden, Stanworth and Chang (2016) found no significant disadvantage between gamers and non-gamers in virtual environments when following instructions. However, Murias, Kwok, Castillejo, Liu and Iaria (2016) show that previous experience of playing video games with a navigational component leads to a higher dexterity with game controls, when comparing with less experienced individuals. Experienced players use more efficient strategies regarding navigation, more specifically, adopting procedural approaches or using cognitive maps.

Purpose and research questions

The purpose of this study will be to examine different types of instruction methods for a virtual reality (VR) environment, and how previous experience may affect the performance. This will be examined through instructions given in text or visual objects. The text instructions are short directions in a limited space in the environment, and the visual objects are objects that gives the user an idea of the tasks goal. The task is to build a tower of cubes in a certain color scheme as fast and correct as possible, where one correct placed cube generates one score.

The research questions for the study are formulated as follows:

1) In which condition do the participants complete the task with the highest score for a certain period of time?
2) In which condition do the participants evaluate the task as easiest to complete?
3) How does previous gaming experience, in either virtual environments and/or computer games, affect performance on the task?
4) What difficulties do the participants experience when interacting in a new virtual environment?

Hypothesis

We believe that the user will generate higher score in task performance in the condition with visual objects, and that the user will evaluate this condition as the easiest to follow, and that previous experience in gaming can affect the performance. Also, we believe it to be important to grasp if the participants experience any difficulties with the environment.

Theoretical background

Several theories explain the difference between modalities and the processes behind them. According to Ganier (2002) information is processed differently when presented as text instead of pictures, since pictures, more directly, lead to constructions of mental representations.

Executive functions such as working memory may play an important role. The working-memory works in such a way that it contains only the most recently activated part, or the conscious part, of the long-term memory, and moves this information to a temporary memory storage (Dosher, 2003, cited in Sternberg & Sternberg, 2012). Earlier studies exploring working-memory for pictures show that pictures are easier to detect compared to printed words (Potter, 1976).

Another study performed by Gozli, Bavelier and Pratt (2014) investigated sensorimotor learning and found that participants with more video game experience increased their sensorimotor learning faster than less experienced participants. However, no significant difference was found in the initial stages of the experiment.

Method

Design

This study examined different types of instruction methods and how previous experience may affect the performance, through user tests. A between-group-design was used, where half of the participants received instructions through text, and the other half received instructions through visual objects. The results were later compared and analyzed. The evaluation was done both quantitatively and qualitatively, through observations, the participants' scores and a questionnaire. The main reason for combining quantitatively and qualitatively data, was to strengthen the interpretation of the results.

Participants

Participants were gathered via convenience sampling to execute tests in a virtual reality environment. There were 34 participants in the study, 14 females, 19 males and one gender fluid. The age had a range of 19-31 years old ($\bar{x}$=22.9).

Procedure

The study took place in a quiet room inside the E-House at Linköping University. Before the tests started the participants signed an informed consent and executed a color blindness test. If the participant proved to color blind they were discarded from the study, if not, they could proceed to the test. Any participant with reading disabilities were assigned to the visual object condition. The rest of the participants were pseudo randomly assigned to one of the conditions.

In the beginning of the test the participant was placed on a marked spot on the floor and the experimenter explained how the controls worked before the participant put on the head-mounted display. To avoid a major difference between those who had tried VR before, and those who had not, all the participants started with a short familiarization phase where they could familiarize with the environment and the controls. The familiarization scene was a white room with 32 cubes; 8 blue, 8 red, 8 yellow and 8 green. The cubes were placed in separate piles

according to colors and in front of the cubes was a pink square, on which a tower could be built. Above the square were invisible containers which lighted up when a cube was dragged over them. This made sure that the cubes would get sucked into the right place, and not fall over when building a tower. The users were encouraged to try and place cubes there to practice before the actual test. Before the test began the participants were given instructions on how the test would proceed and asked to stand on the black square on the floor, facing the direction of the pink square.

When the main test started the user was placed in a similar room as in the familiarization phase, with instructions on how the tower was to be built placed next to the pink square. In the condition with visual objects there was a semi transparent tower with different colored cubes (see Image 1), and in the visual text condition the tower was exchanged for printed text (see Image 2), in the same place as the cubes. The order of the colors were randomized for each room. The participants had 27 seconds to build a tower with as many correct cubes as possible, and the experimenter counted down from three before starting the trial. After 27 seconds the participants were automatically moved to a pause-scene which looked like an empty white room. After a few seconds, with room for possible questions, the participants were transferred to next trial. In this trial the room looked exactly like the previous room, only with a different order on the colors in the instructions. All in all, the participants repeated the test in 7 different rooms. After this the participants completed a self evaluation form and were then debriefed.

Recordings were saved and analyzed together, after all experiments had been executed. The possibility to rewind sequences of the video was enabled. The top 2 performers and the bottom 2 performers of each condition were thoroughly analyzed. These participants were chosen based on their total end score.

Image 1. Illustrates the visual object condition, where the instructions were in the form of a tower of cubes in a random color scheme.

Image 2. Illustrates the visual text condition, where the instructions were in the form of a tower of text in a random order.

Pilot Study

A pilot study was conducted to identify possible weaknesses in our user tests. We decided on that 27 seconds was a good time for the test and that 7 different trials were enough for the participants to perform.

Results

Quantitative results

Table 1 illustrates that the text condition had a higher mean in all rooms except in room 5 where text had a mean of 7.06 and visual objects had a mean of 7.82. However, the text condition had a higher standard deviation than the visual object condition in all but two rooms: room 5 and room 7. The text condition had its lowest

mean in room 1 with 5.82 correct cubes and the highest mean in room 4, with 8.18 correct cubes. The last room had a mean of 8.06. The object condition had its lowest mean in room 1 with 4.53 cubes and its highest in room 5, with 7.82 cubes. The last room had a mean of 7.71 correctly placed cubes. An independent T-test exposed that the only room with $p<0.05$ was room 1, and the lack of significance in the other trials indicates that the instruction types did not affect the participants' scores.

Table 1. *Mean values and Standard Deviation for the text condition and the visual object condition*

Condition	Calculations	Room1	Room2	Room3	Room4	Room5	Room6	Room7
Text	Mean	5,82	6,41	6,47	8,18	7,06	7,59	8,06
	Std. Deviation	1,976	1,839	1,807	2,378	1,56	2,526	2,487
Visual Object	Mean	4,53	6,06	5,88	7,12	7,82	7,06	7,71
	Std. Deviation	1,328	1,088	1,867	1,495	1,704	1,819	2,285

Regarding gaming experience the result showed that there were 10 participants that played a higher amount of video and/or computer games in the text condition and 6 participants in the visual object condition. When comparing people who played a lot of videogames and those who did not the participants were divided into two different groups according to their answers in the questionnaire. The questionnaire contained a question on the familiarity with video and computer games where the participant was to score from one to six how much they considered themselves to play video and/or computer games. Those who had answered on the three lowest levels of computer gaming were paired into one group, and will henceforth be called non-gamers. Those who had answered on one of the three highest levels were paired into another group named gamers. The difference between these two groups was small. They all showed a development over time, but non-gamers had their peak in the end, with room 4 only slightly behind room 7. Non-gamers had a small peak in room 4, whilst the gamers had a peak in room 5. There were no significant difference between the groups which indicates that the familiarity with video and/or computer games did not affect how well the participants performed in the tests.

Qualitative results

The open-ended questions were analyzed to find recurrent themes. The three themes that finally emerged from the analysis were: natural for being a virtual environment, easy to get used to and unnatural with laser pointer.

The observation revealed three noteworthy findings, namely precision, motor skills and hand techniques. It might be important to note here that all these three findings can be said to play an important role in a virtual environment, and that they either directly or indirectly may affect each other.

Firstly, top performers, in general, had a high accuracy with their controls. A higher level of precision tended to lead to less time needed to correct mistakes, and thereby a higher end score. It seems reasonable that precision may have affected the participants' final scores.

Tightly connected, though different, is the participant's motor skills within the world. Participants who could navigate well to surprising events within the virtual world showed less difficulties in coping with the main task. Therefore, similar to precision, a higher level of motor skill should facilitate a participant to reach higher scores.

Another discovered finding was the difference between gamers and non-gamers

when it came to hand techniques. Both participants who used their hands simultaneously and alternately, are included in the two-handed strategy. The one handed strategy implies that the participant used only one hand to move the cubes. The result showed that gamers used more two-handed strategies compared to non-gamers. High performers, however, showed a tendency to successfully apply an alternating two handed-technique where they alternated between an occupied and an unoccupied hand. This resulted in a higher efficiency which in the end resulted in a better end score. However, not all high performers did this, and some participants that used the two-handed technique received a low score. There were also examples of high performers using a one handed technique.

Discussion

Discussion of results
This study has investigated several factors that may affect performance in VR. However, neither of the instruction methods nor the previous experience were shown to be significant in any of our tests. There may be several reasons why our results conflicts with previous research such as Ganier (2002), Potter (1976) and Healy et al. (2016). A reason may be that the overall task was not complex enough to produce a difference. With more thorough instructions for a more complex task, the results may have been different and more in line with reality.

There were several cognitive processes that could have been explored. For example, the study did not investigate working-memory and how it may have affected the results. The user may intentionally have tried to remember the instructions in order to become more effective, and in that way finish the task faster. Instead they were moving active information to a temporary storage (Dosher, 2003, cited in Sternberg & Sternberg, 2012). Earlier studies exploring working-memory showed that pictures were easier to detect compared to printed words (Potter, 1976), but unlike that study no significant difference was shown in our study.

Moreover, the results showed that the learning effect for both conditions was not very prominent. However, a clear learning effect could be seen, as participants always produced a better score on the last tower than the first one, no matter which condition they were in. The scores peaked in room 4 for the text condition and room 5 for the visual object condition. After the peaks the score decreased and then next room increased again. The peaks could depend on the color combinations in these rooms being easier to build. A reason for the decreases after the peaks could be that the participants tried new strategies. Findings from the observations revealed that some participants attempted to change their strategy during the test. For example, some participants attempted a two-handed technique from a one-handed technique in the middle of the test, which in some cases led to a local dip in their performance.

The observation indicated that there were different techniques available in order to approach the task. Many high performers, however, mainly used an alternating two handed technique. It could be concluded that this is indeed a successful technique to solve the task. However, as seen by some high performers, it is not necessarily the sole best strategy. A high precision one handed technique can outweigh the advantage given using a two handed technique, as it is possible that a split-attention effect is taking place when using a two handed technique.

There were examples of participants using both these approaches during the experiment that ended on both ends of the result scores. Seemingly, the choice of strategy could be regulated by factors such as participants' sensomotorics and precision. Possibly, using a two handed technique could lead to a split-attention effect which could affect precision. For future studies this would be an important factor to consider within VR. A more exact relationship between two handed use and the degree of precision loss could be an interesting topic to investigate further.

Contrary to our expectations, participants who rated themselves as playing

more games did not significantly outperform the other group. This partly contradicts the findings from Gozli et. al (2014) who found that participants playing video games minimum 3-4 times per week increased their ability to learn sensomotorics over time, and Murias et. al's (2016) study which stated that previous video game experience generated a higher precision with game controls. It is possible that the experience of playing video games on gamepads does not carry over to virtual reality, possibly because VR requires a different type of motor skill.

Discussion of method

When evaluating the method it was apparent that some things could have been done differently for a more certain result. One thing that was discussed was the choice of design. The study was built up by a between-group-design to make sure that there was no learning effect affecting the participants. One advantage with using within-group-design instead is that it cancels individual differences. With this in mind, it was decided that the between-group-design suited this study better, as it was important to avoid any learning effect.

Furthermore, the design of the questionnaire could have contained other types of questions and been structured differently. The results displayed three overall themes; natural for being a virtual environment, easy to get used to, and that he laser pointer felt unnatural. Another structure on the questionnaire could have generated more interesting themes, and less need to rely heavily on observations and scores to detect differences between the conditions. With the facts on hand we could have included more questions about the instructions to generate more interesting patterns. However, in hindsight, a within-group design would have been more appropriate.

Conclusion

This study has investigated instruction types and gaming experience in VR. It found no significant difference between the two conditions of text and visual objects, and neither any difference in the end scores depending on participants' gaming experience. However, other factors may have affected these results. Something the study found was that many participants who rated themselves as playing more games tended to use the controls more swiftly, and often used two handed techniques. Moreover, for some participants one handed strategies resulted in higher precision and hence fewer mistakes. This study has highlighted several areas within virtual reality that would be interesting for further investigation.

References

Bates-Brkljac, N. (2012). *Virtual Reality: Computer Science, Technology and Applications*. New York: Nova Science Publishers.

Chen, C. J. (2006). The Design, Development and Evaluation of a Virtual Reality Based Learning Environment. *Australasian Journal Of Educational Technology*, 22(1). doi:10.14742/ajet.1306

Ganier, F. (2002). Processing text and pictures in procedural instructions. *Information Design Journal*, *10*(2), 146-153. doi:10.1075/idj.10.2.12gan

Gozli, D., Bavelier, D., & Pratt, J. (2014) The effect of action video game playing on sensorimotor learning: Evidence from a movement tracking task. *Human Movement Science, 38,* 152-162. doi:10.1016/j.humov.2014.09.004

Healy, A., Schneider, V., McCormick, B., Fierman, D., Buck-Gengler, C., & Barshi, I. (2013). Which modality is best for presenting navigation instructions? *Journal Of Applied Research In Memory And Cognition, 2*(3), 192-199. doi:10.1016/j.jarmac.2013.07.004

Murias, K., Kwok, K., Castillejo, A., Liu, I., & Iaria, G. (2016). The effects of video game use on performance in a virtual navigation task. *Computers In Human Behavior, 58,* 398-406. doi:10.1016/j.chb.2016.01.020

Potter, M. (1976). Short-term conceptual memory for pictures. *Journal Of Experimental Psychology: Human Learning & Memory, 2*(5), 509-522. doi:10.1037//0278-7393.2.5.509

Sternberg, R. and Sternberg, K. (2012). *Cognition*. 6th ed. Belmont: Wadsworth, cop. 2012.

Warden, C., Stanworth, J., & Chang, C. (2016). Leveling up: Are non-gamers and women disadvantaged in a virtual world classroom? *Computers In Human Behavior, 65,* 210-219. doi:10.1016/j.chb.2016.07.033