Bachelor of Science in Digital Game Development
June 2022

Assessing the affect on short-term memory in students by comparing a serial recall Augmented Reality game and a card version.

Linus Dorell
Oskar Nyman

Faculty of Computing, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden

This thesis is submitted to the Faculty of Computing at Blekinge Institute of Technology in partial fulfilment of the requirements for the degree of Bachelor of Science in Digital Game Development. The thesis is equivalent to 10 weeks of full time studies.

The authors declare that they are the sole authors of this thesis and that they have not used any sources other than those listed in the bibliography and identified as references. They further declare that they have not submitted this thesis at any other institution to obtain a degree.

Contact Information:
Author(s):
Linus Dorell
E-mail: lido18@student.bth.se

Oskar Nyman
E-mail: osny19@student.bth.se

University advisor: Eng. Diego Navarro, Tek.Lic.
Email: dna@bth.se
Department of Computer Science

Faculty of Computing
Blekinge Institute of Technology
SE–371 79 Karlskrona, Sweden

Internet : www.bth.se
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57

Abstract

Background AR technology has been increasing across domains in cognitive activities, i.e. learning. Although there are studies that try to examine the quality AR can bring into different territories of human culture, such as educational settings, few studies aspire to determine how AR can have on human memory. Particularly short-term memory.
Objectives This research aims to assess the affect on short-term memory through an AR and Analog game.
Methods The method proposed for this thesis work is a user study in a controlled environment to gather data for the results. In order to test the hypothesis, a quantitative approach was selected as two versions of the same game were compared. A within-participant experiment was designed.
Results The results from the experiment indicate that AR has a lower score on average compared to its non-virtual counterpart.
Conclusion Overall, our findings suggest that AR does not have a significant affect on short-term memory with digits.

Keywords: Augmented Reality; Short-term memory; Computer-aided instructions; Webcam application; Mixed Reality.

Acknowledgments

We thank Diego Navarro, our supervisor, for giving us constant feedback and advice for the thesis and giving us access to an experiment room. We would also like to thank BTH IT Help-Desk for letting us use their equipment to set up the AR game along with access to university computers. Lastly, we want to thank all the participants for allowing us to collect data for the thesis.

Contents

Abstract i

Acknowledgments iii

1 Introduction 1
- 1.1 Aim and Objectives . . . 1
- 1.2 Research Question . . . 2
- 1.3 Expected Outcomes . . . 2
- 1.4 Background . . . 3
- 1.5 Document Layout . . . 3
  - 1.5.1 Introduction . . . 3
  - 1.5.2 Background . . . 3
  - 1.5.3 Related Work . . . 3
  - 1.5.4 Method . . . 4
  - 1.5.5 Results and Statistical Analysis . . . 4
  - 1.5.6 Discussion and Analysis . . . 4
  - 1.5.7 Conclusions and Future work . . . 4

2 Related Work 5

3 Method 7
- 3.1 Participants . . . 7
- 3.2 Implementation . . . 7
  - 3.2.1 Card Version . . . 7
  - 3.2.2 AR Version . . . 8
- 3.3 Gameplay . . . 9
- 3.4 Scoring System . . . 10
- 3.5 Experiment Procedure . . . 10
  - 3.5.1 Pilot Experiment . . . 10
- 3.6 Statistics . . . 11
- 3.7 Ethics . . . 12

4 Results and Statistical Analysis 13
- 4.1 Five digit sequence . . . 13
- 4.2 Six digit sequence . . . 14
- 4.3 Seven digit sequence . . . 14
- 4.4 Eight digit sequence . . . 15
- 4.5 Nine digit sequence . . . 16

4.6 ANOVA assumptions 17
4.7 Wilcoxon nonparametric test 18
4.8 Paired t-test 18

5 Discussion and Analysis **19**
5.1 Validity threats 19
5.2 Limitations to the study 20
5.3 Significance of the study 21

6 Conclusions and Future Work **23**
6.1 Conclusion 23
6.2 Future work on research subject 23

References **25**

A Supplemental Information **27**

A Consent Form **29**

List of Figures

3.1 This figure shows an example of what a participant has selected. The top part of the figure represents how the digit sequence is hidden underneath a cover. 8

4.1 Five digit sequence: Rounds One and Two. 13
4.2 Six digit sequence: Rounds Three and Four. 14
4.3 Seven digit sequence: Rounds Five and Six. 15
4.4 Eight digit sequence: Rounds Seven and Eight. 15
4.5 Nine digit sequence: Rounds Nine and Ten. 16

Chapter 1

Introduction

Techniques, innovations, technology and research studies have been conducted to assess effects on memory [18]. Previous research suggests an effect on short-term memory by using Augmented Reality (AR) [8], AR is a way to see and experience the real world with the aid and help of computer-generated. However, according to Wu (2008) "Klopfer(2008) [6] indicated that the term AR should not be defined restrictively. This term could be applied to any technology that blends real and virtual information in a meaningful way." [17]. In a study by Klopfer and Squire (2008), "AR could be broadly defined as "a situation in which a real world context is dynamically overlaid with coherent location or context sensitive virtual information""(p. 205) [7].

Some studies research how AR can improve short-term(ST) memory recall. A study by Sommerauer and Mueller in 2018 concluded that short-term memory performance with AR was better in the short run but not in the long run [16]. Another study conducted by Chen et.al in 2019 , through a memory recall experiment found that the non-virtual test gave overall better results, while the AR made the participants feel less anxiety and more committed [3]. This study assessed how student's short-term memory is affected by comparing score differences when using two serial recall games. The first serial recall game version used AR on a webcam device while the second version used printed cards. According to Roediger et.al, serial recall is the process of presenting a randomly ordered sequence of numbers, letters or images to a user, who is then required to repeat them in the same order [12] [17].

Universities and institutions have come to divergent conclusions concerning how AR can be used for educational purposes. The conflicting results show that there is an inconsistency in how effective AR is as a medium in the educational setting [16] [3] [5] [13]. Therefore, the objective of our thesis is to research and assess if there is an affect on memory through the lens of AR. The scope does not allow for a thorough AR memory application to be created for this degree project. Instead, a serial recall game will be used to test an individual short-term memory.

1.1 Aim and Objectives

The study aimed to assess the effect on short-term memory in students by testing a serial recall AR game when compared to a card version. To reach the goal, 21 participants partook in a within-subject experiment and were exposed to both versions of the serial recall game. The participants were tasked to remember sequences of digits

in ascending level of difficulty and then attempt to recall the sequences. The data collected from the experiment was categorised and a score was calculated based on two parameters, digit ordering and correct digit. The data was then analysed and compared to discuss the findings and relate to previous research and draw a conclusion. After the hypothesis was tested and the thesis is completed the data collected can be deleted.

The following objectives were completed in order to reach all of the previously stated goals.

1. Design the AR and card game.
2. Develop the AR and card game.
3. Design the experiment.
4. Conduct the experiment.
5. Collect data from the experiment.
6. Analyse the data and compare the results.
7. Discussion of results and previous research as well as hypotheses testing.
8. Develop a conclusion.
9. Future work.
10. Delete the data collected after the thesis dissertation has been approved.

1.2 Research Question

How can students short-term memory be affected in a serial recall augmented reality game, when compared to a card version?

The proposed hypothesis was that an AR game will affect the short-term memory in students by improving their ability to recall longer sequences of numerical digits.

1.3 Expected Outcomes

The main expected outcome of the study was that the score gathered from the two games would show a substantial difference, presenting a better performance when recalling digits with the AR game than in the non-virtual version.

The project delivered two games, both of which presented the user with a set of digits. The AR version used 3D assets to represent the digits, while the non-virtual game-used cards had digits printed on them. Another expectation was the thesis dissertation itself and the results of the study.

The proposed methodology generated results based on player performance which was then compared between paired sample groups. The sample data was analysed to determine a statistical significance by using a paired t-test or a Wilcoxon signed-rank test. Finally, the hypothesis would be rejected or validated by discussing and motivating the differences in the results.

1.4 Background

Memory is the act of storing a certain amount of data and using the data later, in 2016 Sherwood states that "Memory is the storage of acquired knowledge for later recall" [15]. Short-term memory is the ability to hold information for a short period of time, Öğmen stated in 2016 that: "STM is severely limited in capacity and can hold information for several seconds to minutes." [19]. STM stands for Short-term memory. The average human can allocate between five to nine objects in their short-term memory, occasionally referred to as Miller's law which suggests there is a 7 ± 2 span of objects held in the memory [9]. This is critical information for how the design of the serial recall game will be implemented.

AR is a way to see and experience the real world with the aid and help of computer-generated perceptual information. In 2013 Wen-Yu stated that: "AR can be defined as a system that fulfils three basic features: a combination of real and virtual worlds, real-time interaction, and accurate 3D registration of virtual and real objects." [17]. While the earliest AR prototype was introduced in the early 1990s for the U.S Air forces [14], it has spread to the gaming, entertainment, education, communications and medicine business. [10]. This will be further explained in related work 2. As previously mentioned in the introduction, AR is a way to experience the real world with the aid of digital models or virtual information. This implies that AR is not tied to a specific platform such as a mobile device. Therefore, by using a webcam for this study, AR can still be used as a stimuli for the experiment.

1.5 Document Layout

1.5.1 Introduction

A brief introduction to what the thesis is about and the area of research.

1.5.2 Background

This is were relevant background information is presented, with brief explanations of Augmented reality and Short-term memory.

1.5.3 Related Work

This is where we present the related work the thesis.

1.5.4 Method

This is were the methodology is presented along with participant criteria, implementation, scoring, choice of statistical analysis and ethics.

1.5.5 Results and Statistical Analysis

This section will present the results from the experiment and the statistical analysis that were conducted along with their results.

1.5.6 Discussion and Analysis

In this section an analysis of the results will be presented, furthermore the results will be discussed and put into a wider context.

1.5.7 Conclusions and Future work

The final chapter will present the thesis conclusions and potential future work.

Chapter 2

Related Work

As of 2022, there have been attempts and research on the development of AR applications. Several institutions and organisations are currently working in that area and focusing on different aspects significant and relevant to current society.

The University of Washington developed an AR application on mathematical concepts for educational purposes for children. The study conducted a user study at a local children's museum, having a total of 27 participants. Kang stated in 2020 that "While ARMath demonstrates the potential of AR for everyday math, more work is needed to address usability issues, design effective child-AI interaction, and evaluate learning." [5].

Another related work of AR comes from the University of North Carolina at chapel hill. Their study involved developing an AR application-specific to head and neck surgery. They used a face validity and case study to evaluate the application made. The results found that AR could be used in otolaryngology–head and neck surgery with current modern technology, improving surgical and patient safety [13].

In 2009, Papagiannis at York University developed an AR experience focused on the field of visual art. Through theory and practical experiments (Hockney's Film Experiment), it came to the conclusion that AR can be used as a medium for a form of cinematic experience [11].

The university of Lisabon developed an AR tool to visualise solar radiation data on facades on buildings. An empirical user study with 22 participants was conducted with an AR application on a Asus tablet. The results of the study present a favourable conclusion about AR. Carmo 2014 stated that "We presented an AR application for the visualization of solar radiation data on the facades of buildings and conducted a user study to assess the usability of the application, which showed an overall favourable opinion of the participants." [2].

A study by Moro in 2017 shows how AR could be used in medical- and health sciences stating: "Both VR and AR are as valuable for teaching anatomy as tablet devices, but also promote intrinsic benefits such as increased learner immersion and engagement" [10]. A large number of research articles, journals and literature exist describing how such training can be used to measure the effect on memory span. According to Korkelia in 2021: "Today, extended reality (XR) technologies such as

VR (virtual reality) and AR (augmented reality) have been touted in influencing memory performance, especially memory recognition, which is one of the most important areas in memory research". This shows that there is a valuable connection between AR and learning memory that could be further researched upon [18].

The works presented here is to show two things, a problem and a motivation. These related works show that there is a pattern to integrate AR into educational settings. However, with the conflicting conclusions, it will be difficult to assess if AR has a significant affect on memory. So wouldn't it be interesting to asses if there really is an affect on ST memory through the lens of AR? The related works do not delve into statistical measurements of how AR could affect ST memory by a serial recall game. To solve this, the study aims to assess the affect a AR game can have on a student's short-term memory, by comparing the data result from two different serial recall games.

Chapter 3

Method

The method proposed for this thesis work is a user study in a controlled environment. In order to test the hypothesis, a quantitative approach was selected as two versions of the same game were compared. A within-participant experiment was designed. This meant that the participants would test both conditions of the stimuli; AR and card game. The independent variable that would change was the choice of medium, AR or card version. While the dependent variable was the score, which represented a quantitative measurement of the short-term memory. Selecting a within-participants study design meant that it generally requires fewer participants, usually around 30 participants would be needed. An issue that arises with the within-participants design, however, is that there is a high risk of carry-over effects between the two variables. In order to avoid this a counter-balancing measure was designed in which the participants would alternate which version of the game they would begin with.

3.1 Participants

The participants for the study needed to fulfil specific criteria. According to a study in 2015 written by Hartshorne, ST memory peaks in a humans early 20s [4]. Therefore the students were between 18 and 35 years of age to yield reliable measurements. The study however decided to use the 18-35 age range to have the possibility to include more participants in the study.

The ongoing Covid infection spread could affect the availability of participants. Mitigation was to advise that participants are vaccinated and healthy. Another option was to disinfect and clean equipment between the participant testing.

3.2 Implementation

Two versions of the same game were created for the experiment, an analog card game and a digital AR game. The following section will describe both versions and their respective rules, game logic, and creation.

3.2.1 Card Version

The card version was created before the AR version which would enable early testing as it required less time to implement. The cards were made in Adobe Photoshop

and later printed at BTH. Four copies of each digit ranging from zero to nine were printed on separate pieces of paper which were used by the participant to select the desired digit combination. Four A4 papers were used for this as all nine digits could fit and then be cut out, the outcome is shown in 3.1.

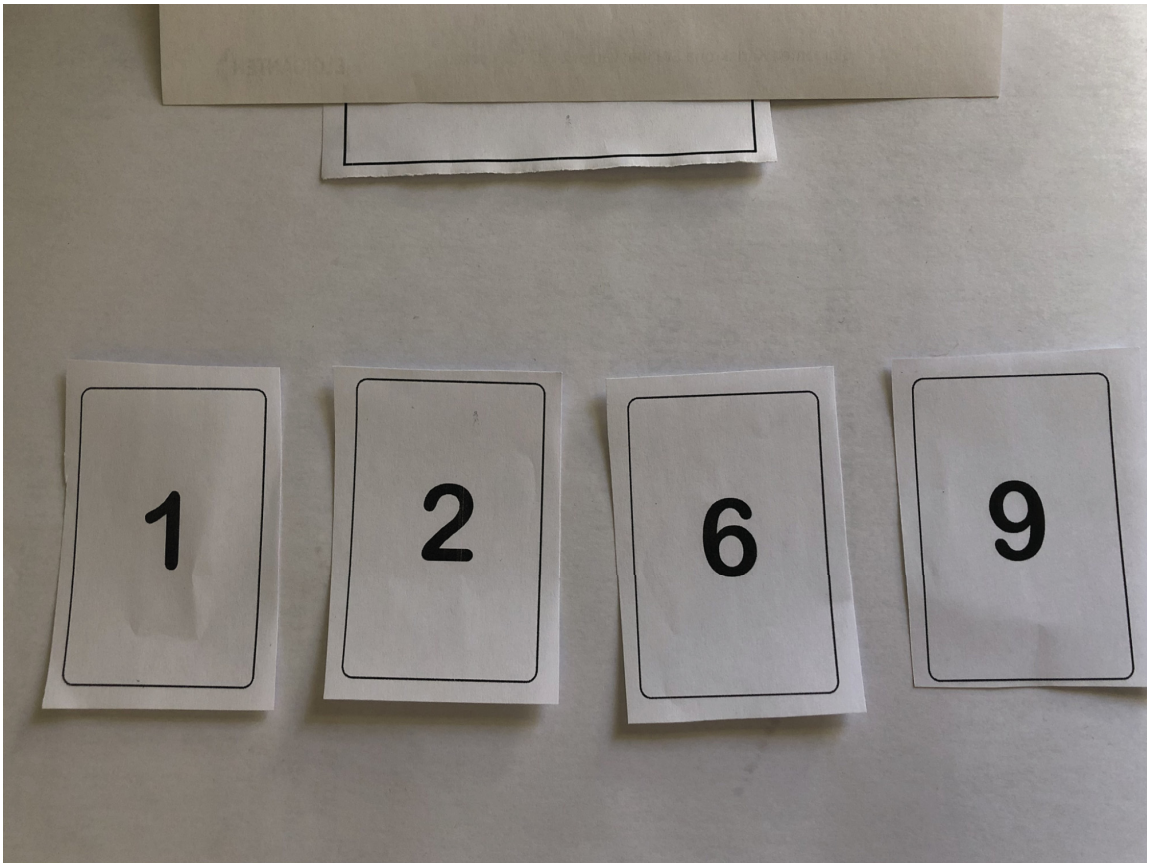


Figure 3.1: This figure shows an example of what a participant has selected. The top part of the figure represents how the digit sequence is hidden underneath a cover.

The ten sequences of digits were printed on single pieces of A4 paper. The digits were printed with black ink while the background colour was white. A black border was also present on all cards. During the game the pile of digit sequences were turned face down so the participant wouldn't be able to see the upcoming round. However, due to the A4 papers being white it was possible to see through the back of the paper. In order to cover the pile of digit sequences two blank A4 papers were stapled together so they could be used as a way to hide upcoming sequences.

3.2.2 AR Version

The AR application was created in Unity using a plugin lib called Vuforia Engine [1]. Vuforia was used as a solution to recognise, track and augment the image through a web-camera. The numbers were placed in the center of the image, allowing for a clear unobstructed view. The digits had a red color and placed in front of a black background in order make them more readable. A digit layout was created to allow

the input of numbers, which was saved in a text file. The contents of the text file was, after the game, copied and saved on Excel after which the text file was deleted. A scene (level) manager was created to keep track of when scenes needed to change. To keep track of the digits and current rounds a round keeper script was created. The timer could be altered depending if a certain round was reached.

3.3 Gameplay

The participants attempted to recall a series of digits in the correct order as the digit sequence. To collect reliable results, the game was designed to stress the short-term memory through a slow buildup in difficulty spread throughout ten rounds. The number of digits that appeared was between five and nine depending on the level of difficulty. The participants were awarded score results based on how accurate they could recall the digits that appeared for a certain amount of seconds depending on the round.

The games were played in two different versions with the same gameplay. The player was presented with a sequence of digits ranging from five to nine. The number of digits in the sequence depended on the number of rounds played, there was a total of ten rounds per version. When the first sequence of digits was presented the player had five seconds to memorise the digits. Once the time limit was reached the player was presented with an interactive part. The player had to recall the order of the sequence and the correct digits by selecting digits from an interface. The chosen digits could be retracted and replaced if so desired with no time limit.

The analog card game had a digit selection that involved the player hand selecting from piles of printed cards with digits. The player selected and placed the individual cards on a table in front of them to form a sequence. Thereafter, the selection was entered into an Excel spreadsheet by the researchers, and the player returned the digits to their respective piles. On the other hand, the AR version presented the player with a keypad, in which the player could use the mouse to select the digits. The digits selected were displayed on the screen and could be removed by pressing a button named 'BackSpace'.

Once the player had finished the interactive stage the game continued to the next round. Depending on the round there was a different time limit and amount of digits per sequence. Every other round the number of digits in the number sequences was increased by one. The time limit for the first four rounds was five seconds, round six and seven had a time limit of seven seconds while for the last four rounds the time limit was nine seconds.

When the player had completed all ten rounds a five-minute break ensued. After the break, the player returned and began the next part, which was the other game version. As explained above, this was alternated between each participant in order to avoid carry-over effects.

3.4 Scoring System

As this study employed a serial recall game the ordering of digits and the correct digits were important. This meant that a scoring system was build based on those two parameters. Therefore, the data collected was represented with a score that was calculated depending on the correct digits recalled as well as the correct ordering. One point was given if the correct digit was selected, one point was awarded for the correct ordering of a digit. The maximum amount of points for a sequence was $n * 2$ in which n is the number of digits in the sequence. The calculation of the score was done manually by the researchers after all the experiments were completed, and the score was double checked.

3.5 Experiment Procedure

The experiment took place in a lab room at Blekinge Institute of Technology (BTH). This section describes and lists the stages that was needed to gather data from the experiments in which item one is the first stage of the experiment.

1. Pre-instruction stage: A small briefing of the experiment and consent form signing.

2. Interface Training: Explanation of rules for the AR and non virtual game both in written and verbal.

3. Experiment: Conduction of the experiment.

4. Results: Collect and enter results to Google Sheets.

Participation during the experimentation was voluntary, and participants had the right to withdraw without giving a reason. Furthermore, a consent form A was given to the participants for careful study and signing, confirming that the participants had taken part and understood the information. The consent form also informed the students who conducted the survey and the supervisor, along with the appropriate contact information.

3.5.1 Pilot Experiment

Prior to the experiments, the researchers and supervisor conducted a pilot experiment in order to evaluate the execution and style of the experiment. As a conclusion of the pilot experiment there were several points that needed to be addressed.

A major aspect was the need for a calm and silent environment which can otherwise lead to detrimental results. This could potentially introduce errors in the results as there would be outside factors that interfered with the experiment and stimuli.

The card version suffered from minor drawbacks such as the cards with the sequences were too small and the participant had to physically bend closer in order

to read. This was solved by making the digits fit on an entire A4 paper, which in turn made it easier to be laid out in front of the participant. Another drawback was that the card selection piles were not ordered correctly but rather randomised. This would add yet another outside factor that could harm the testing of the stimuli and was not intentional. To resolve the issue the cards were laid out in the correct order from left to right, ascending. As there was not a lot of space on the table the first five digits were placed in one row while the rest of the digits were placed in a row below.

The AR version had substantial issues that needed to be improved upon. One such issue was that the 3D models for the digits had different designs which could lead to participants spending more time trying to discern which digits were displayed. Another issue was related to the background of the display, in which it was possible to see the real world behind the sequence of digits. This was considered to be too distracting and could lead to faulty results, therefore it was decided to add a plane in the game that would cover the entire background. This solved an issue regarding the digits sequences not having enough contrast to the background. As the plane for the background was black the digits were coloured red in order to create a strong contrast in colours. Another issue was that the digits were skewed and didn't fit in the target area which meant that the beginning and end of sequences were cut off. This was resolved by moving the web-camera further back so that the sequences fit without issue on the screen. The final problem that was discovered was that the web-camera couldn't detect the target image in some cases. Therefore a A4 paper with a clear logo was printed out and placed on a whiteboard in order for the web-camera to detect the logo easier.

3.6 Statistics

The results gathered from the experiments and the dependant variable, score, would show a behaviour between the two independent variables. To analyse the results and therefore test the hypothesis a paired t-test analysis was suggested, as long as the ANOVA assumptions are met. Otherwise, other statistical methods, such as the Wilcoxon signed-rank test, would be used.

A paired t-test is used to compare two population means in which observations in one sample can be paired with observations in the other sample. This type of test is most commonly used to measure studies of repeated measures or case-control studies. Since this study had one factor (the game) and two levels (card and AR), a paired-sample t-test was appropriate.

On the other hand, if the ANOVA assumptions were not met a Wilcoxon test would be used. The Wilcoxon signed-rank test is a non-parametric statistics test that compares two paired or matched groups. It is the non-parametric equivalent of the paired t-test. It is used to test for a significant difference in the mean(median) of paired observations or measurements. Unlike a t-test, it doesn't require a normality test to be legitimate. A Wilcoxon test is less effective compared to a t-test with its

95% accuracy, however, it is the preferred type of test when it is not possible to use a paired t-test.

As the study intended to use a paired t-test to evaluate if there was a significant difference, there were two assumptions that needed to be fulfilled, a normality (D'Agostino-Pearson) and a homoscedasticity (Levene) test. These assumptions ensure that the data has a normal distribution and the two paired groups have the same variance. The normality test was conducted using Prism GraphPad which lets the user choose between several different versions of the same test. The D'Agostino-Pearson (omnibus K2) test was selected due to it being recommended by the software. A normality test checks if a P value is higher than the threshold of *0.05* in which case it passes the test. The homoscedasticity test was conducted in Excel using the Levene test. The Levene homoscedasticity test was selected due to it being recommended by the supervisor of this study. Similarly to the normality test, a P value higher than *0.05* is considered a pass. In chapter 4 the pre-requisites tests will not show the P values, only if a round has passed or failed will be displayed.

3.7 Ethics

As the study gathered potentially sensitive data concerning the participants there were ethical questions that needed to be addressed. The performance data measured the affect on participant's short-term memory can be considered sensitive. The data collected was therefore anonymous to protect the integrity of the participants. All of the participants involved in the experiment had the right to withdraw.

The data collected remained entirely anonymous as personal information regarding participants was not recorded or published. The data from the study was stored and encrypted on the researchers personal computers along with Google Sheets for transfer of data, access to the data was limited to the researches and supervisor. As Google Sheets was used to store the data a copy is also owned by Google.

Access to the data gathered from the experiment was off-limits to anyone that was not a member of the project. There was not a direct link between individual participants and the result data, therefore the project members could not make a connection between the two. After the project achieved its purpose, the gathered data shall be wiped.

Chapter 4

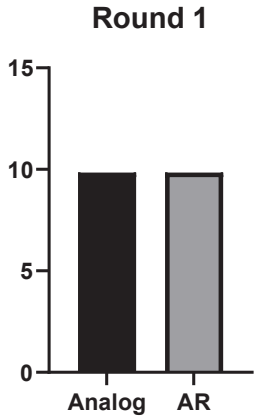
Results and Statistical Analysis

This section shows the results from the experiment. A detailed description on how the statistical tests were performed can be found in chapter 3. Firstly the results from the two groups and the mean of the combined participant scores in each group and their differences will be shown. Thereafter the outcome of the prerequisites will be presented, indicating which rounds will use the Wilcoxon signed-rank test or the paired t-test. Lastly, the Wilcoxon signed-rank tests and the paired t-test will be presented to show which rounds had a significant difference.

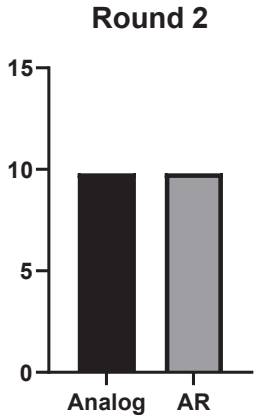
The raw data from the experiment has graphs dedicated to show the mean of the combined participant score. The X-axis on the graphs represents the two versions of the serial recall game. The black bar is Analog and the grey bar is AR. The Y axis represents the mean score for both versions.

4.1 Five digit sequence

The results for round one and two did not have a mean difference between the score of AR and analog.



(a) This figure shows no mean difference between the groups.

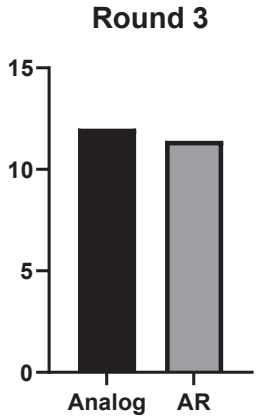


(b) This figure shows no mean difference between the groups.

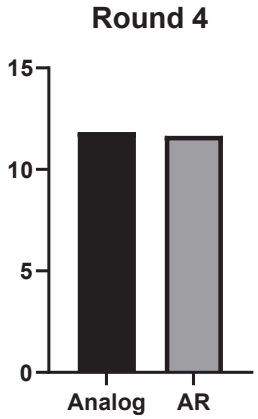
Figure 4.1: Five digit sequence: Rounds One and Two.

4.2 Six digit sequence

The results from round three shows a slight difference between AR and analog. Round four does not show a significant mean difference.



(a) This figure shows a mean difference between the groups, in favor of the Analog version.

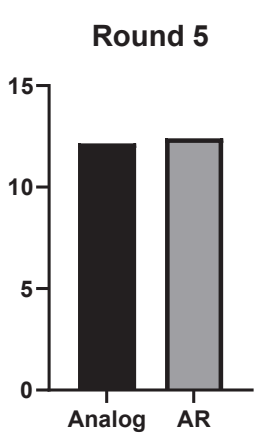


(b) This figure shows no mean difference between the groups.

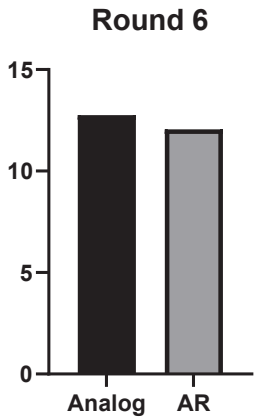
Figure 4.2: Six digit sequence: Rounds Three and Four.

4.3 Seven digit sequence

The results for round five and six shows a slight difference in means, for round five it is in favor of the analog version whereas for round six it is the opposite.



(a) This figure shows a slight mean difference between the groups.

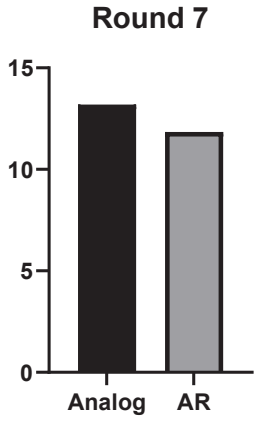


(b) This figure shows a mean difference between the groups, in favor of the analog version.

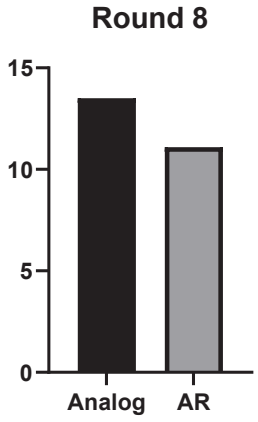
Figure 4.3: Seven digit sequence: Rounds Five and Six.

4.4 Eight digit sequence

The results for round seven and eight shows a mean difference in favor of the analog version.



(a) This figure shows a mean difference between the groups.

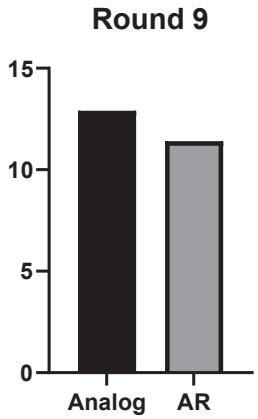


(b) This figure shows a substantial mean difference between the groups.

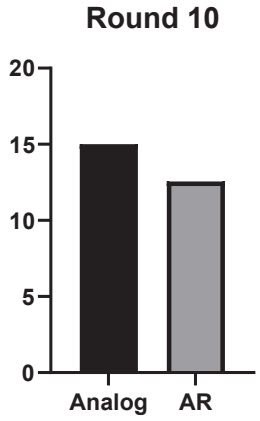
Figure 4.4: Eight digit sequence: Rounds Seven and Eight.

4.5 Nine digit sequence

The results show that for round nine and ten there is a mean difference between the groups, in favor of the analog version.



(a) This figure shows a mean difference.



(b) This figure shows a mean difference between the groups.

Figure 4.5: Nine digit sequence: Rounds Nine and Ten.

4.6 ANOVA assumptions

Table 4.1 shows the outcome after conducting both a normality test and a homoscedasticity test for all the rounds. As described in chapter 3 these tests are conducted in order to determine which statistical analysis is best suited for a certain data-set.

Rounds	**Normality**	**Homoscedasticity**
Round 1	**Analog:** Fail, **AR:** Fail	Fail
Round 2	**Analog:** Fail, **AR:** Fail	Fail
Round 3	**Analog:** Fail, **AR:** Fail	Fail
Round 4	**Analog:** Fail, **AR:** Fail	Fail
Round 5	**Analog:** Pass, **AR:** Fail	Fail
Round 6	**Analog:** Fail, **AR:** Pass	Fail
Round 7	**Analog:** Pass, **AR:** Pass	Pass
Round 8	**Analog**: Pass, **AR:** Fail	Fail
Round 9	**Analog**: Fail, **AR:** Pass	Fail
Round 10	**Analog:** Pass, **AR:** Pass	Pass

Table 4.1: A fail means that the p-value was lower than the acceptable value of 0.05, which suggests that there is no normal distribution in a sample group or an acceptable variance between the two sample groups.

The results after conducting the prerequisite tests show that around seven and ten passed and can therefore be statistically tested using a paired t-test. However, the rounds that failed can still be tested by using a non-parametric test such as the Wilcoxon signed-rank test.

4.7 Wilcoxon nonparametric test

As discussed above in chapter 3, should a sample pair round not meet the requirements of the normality and homoscentricity test, a Wilcoxon test will be used. Based on the previous results, every round apart from rounds seven and ten will use the Wilcoxon signed-rank test.

Rounds	**P $<$ 0.05?**	**P-value**
Round 1	Fail	0.5000
Round 2	Fail	0.5000
Round 3	Pass	0.0313
Round 4	Fail	0.3750
Round 5	Fail	0.3957
Round 6	Fail	0.0742
Round 8	Pass	0.0124
Round 9	Fail	0.0817

Table 4.2: The table above shows which rounds had a significant different p-value, i.e lower than 0.05.

The results of the Wilcoxon test shows that rounds three and eight passed, indicating a significant difference between the sample groups.

4.8 Paired t-test

After having met the assumptions for a paired t-test, rounds seven and ten were statistically compared between the AR and card game. Table 4.3 shows the results of the paired t-test.

Rounds	**Paired t-test with P value**
Round 7	0,0265
Round 10	0,0020

Table 4.3: The table shows P values calculated between the card and AR versions using a one-tailed paired t-test.

The results from the paired t-test shows that both rounds are below the critical value of $p < 0.05$ meaning that there is a statistical significance between the card and AR game versions. The results from the statistical tests will be further discussed in the succeeding chapter.

Chapter 5

Discussion and Analysis

In section 1.2 the hypothesis stated that AR will have an improved affect on a students short-term memory when recalling longer sequences of digits. The statistical analysis of the results shows that 4 out of 10 rounds have a statistical significance, all of which were in favour of the card version. This rejects the previously suggested hypothesis, as AR does not have a significant affect on short-term memory when recalling longer sequences of digits.

Based on the raw data in the results, an overall view of the scores show that the card game version has a better performance rate than the AR counter-part. The earlier rounds, with less digits in the sequences, were relatively similar with the card game being slightly favored. However, the first six rounds (apart from round 3) failed the Wilcoxon test. This means that it is not possible to determine how or what way the performance was affected as it could depend on outside factors. The third round passed with a p value of 0.0313 which contradicts round four and the two rounds before. The reason for this is believed to be due to double digits, which is discussed later in the chapter.

The raw data results for the later rounds showed a higher performance rate for the card version. The paired t-test for rounds seven and ten resulted in a significant difference which suggests that the short-term memory was affected by the choice of medium. Round eight didn't pass the pre-requisite tests and was therefore processed using the Wilcoxon test. The results showed that there was a significant difference, meaning that the choice of media did affect the performance on short-term memory within the conditions of the experiment. Round nine didn't pass the Wilcoxon test but was close to the threshold with a p value of 0.0817. As round nine and ten have the same amount of digits in the sequence it would be beneficial to investigate a possible pattern emergence if more samples were taken.

5.1 Validity threats

This section will outline conflicting data in the results and explain why these conflicts may have occurred. One such conflict was observed for the six digit sequence 4.2 in which round three showed a significant difference in favour of the analog version while round four showed no significant difference. A reason for this conflict could be due to the last three digits in the analog sequence consisted of a double-digit,

two of the same digit, followed by a single digit. This could show that it is easier to recall double-digits in a sequence, compared to two different digits. The pattern reoccurs for the nine digit analog sequence 4.5 in rounds nine and ten, where either one double-digit or two double-digit occurred. As a consequence, this could result in the p values for rounds nine and three being higher or lower than the threshold and thereby jeopardising the validity of the results. To have resolved this issue, more samples could have been taken as well as avoiding the use of double-digits.

There were a couple of conflicts in the results due to the implementation and conduction of the experiment. A conflict arose because the AR version was entirely automated from the start to finish. This gave little space for participant-researcher communication, even though the communication was kept to a bare minimum the participants had fewer interactions with the researchers compared to the card version. This could mean that the participants didn't experience the same amount of stress for both versions, which could skew the results.

The other conflict could be observed in the raw data results where in certain cases the participants didn't enter the correct amount of digits. The rules do not mention that the participants need to enter the correct amount of digits per round. This led to the lack of results in some rounds, especially for the AR version. The reason why it didn't occur as often in the card version was because the researchers told the participant how many digits were missing when asked by the participants. This was a clear difference and had an impact on the results as some of the rounds the participant was rewarded a lower score. The reason for participants not selecting the correct amount of digits seemed to be that they forgot the last digits in the sequences. If this has an active impact on the results is difficult to prove, however, it took away the element of random as the remaining digits were recorded as zero's by the researchers.

Finally, to avoid ordering effects that can occur in these instances, it would have been better to randomize the stimuli instead of alternating for every other participant.

5.2 Limitations to the study

The following section will describe limitations to the study and why they had an impact. One of the major limitations to the study was the participant pool, this was due to the selection process. As the study was to take place at BTH it narrowed the amount of participants down by a substantial amount. If the study could have been conducted in a controlled manner online or included other universities the amount of participants would probably have been higher. However, due to the study being controlled it was difficult to set it up in a way that would meet the requirements.

The device functionality and availability had an impact on the study. The BTH's IT-department did not have the preferred Android devices for the AR application that was developed. Therefore, this study used a webcam instead of a phone camera to mitigate this limitation. This restriction has consequences for the sustainability

and usefulness of the study. The reason for this is because even though AR is not strictly targeted towards a certain platform the use of the webcam directed the study to a more static environment, which may not be an ideal. However, it does not invalidate the findings or results of the study as it represents and tests two different mediums and their effect on ST in students when recalling longer sequences of digits.

As mentioned in the section above, there is a limitation to the results of this study as some digits occurred twice after each other in certain sequences. As these occurrences only took place in the card game there is a clear limit with regards to the validity of the results as there is a bias towards a particular game version. In hindsight, this could have been solved by being more careful of the randomly generated digits.

The final limitation is that the study focuses on digits. The reason digits were used instead of symbols or words is due to its simplicity and nature of using it in a serial recall game. However, the results could perhaps be effected if symbols were used, and it would be interesting to see if there is an effect on the results in a future work.

5.3 Significance of the study

The significance of the statistical analysis from the study shows that in 40% of the rounds there was a significant difference in favour of the Analog version than compared to AR. It could be argued that the impact of the findings may grant a broader and general understanding of how effective AR and Analog could be when it comes to ST in students and recalling longer sequences of numeric digits. Another possible impact could be relevant to how digits and words could be remembered differently if a similar study was conducted with words.

In the section on related work, several universities had already or are currently developing tools concerning AR and how it can be used in education or other fields [2] [5]. Even though this study does not focus on the ability to improve one's ST memory in the long-term, it could be argued that this study sheds light on which medium could be relevant for those that wish to improve their ST memory in the short-term.

This thesis can sustain a discussion on how a serial recall game created using two different mediums, AR and Analog, and the effect they have on a students ST memory. It could also be argued that the results and statistical analysis that are provided in this thesis can help future work around the subject of AR, and how to improve the effect it could have on ST memory. However, on the other hand it may be difficult to justify that the results from this study would have an positive impact on future work's sustainability. A possible consequence of the results found could be that the work done in the future may contradict or outright disprove the results of this experiment. The universities of Washington, York, Lisabon and North Carolina

at chapel hill all worked on some form of AR application at different points of the 2008 - 2020, proving that the research field is still very active [11] [13].

Chapter 6

Conclusions and Future Work

6.1 Conclusion

Based on the data gathered and statistical analysis, the conclusion for the study is that there is not an improved effect on ST memory in students when recalling longer sequences of digits. This rejects the hypothesis suggested in the introduction section 1.2. The raw data from the results show that the greater the length of the sequences of digits a participant needs to remember, the easier it is to see the difference in the score depending on the medium. The statistical analysis did not show a significant difference between the two mediums during the earlier rounds. The later rounds showed an improved ST memory for the analog version based on the statistical analysis. It can be concluded that AR does not offer a significant beneficial effect in ST memory when recalling between seven to nine digits in a sequence. Therefore, one can argue that the traditional analog version is to be preferred when holding between seven and nine digits in the ST memory.

6.2 Future work on research subject

Several potential paths could be further worked upon concerning future work on this research subject. Theoretically, the best possible tests have been applied for each of the data sets. However, increasing the amount of participants would lead to better evidence concerning the effects of the different mediums.

This study shows that the later rounds with more extensive digit sequences impacted the performance depending on the chosen medium. Therefore it would be of additional value to investigate this difference by adding more samples to identify precisely when this difference occurs and how likely it is to happen.

Another future work can involve researching how ST memory can be affected by biological gender and age. This study did not gather that type of data, it is considered an exciting avenue.

Finally, switching from AR ST-memory to VR Spatial memory games is another potential recommendation for future work.

References

[1] (2022) Vuforia developer portal. [Online]. Available: https://developer.vuforia.com/

[2] M. Beatriz Carmo, A. P. Claudio, A. Ferreira, A. P. Afonso, P. Redweik, C. Catita, M. C. Brito, and J. N. Pedrosa, "[poster] visualization of solar radiation data in augmented reality," in *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, pp. 255–256. [Online]. Available: http://ieeexplore.ieee.org/document/6948437/

[3] c. chen, L. Zhang, T. Luczak, E. Smith, and R. F. Burch, "Using microsoft HoloLens to improve memory recall in anatomy and physiology: A pilot study to examine the efficacy of using augmented reality in education," vol. 12, no. 1. [Online]. Available: https://aquila.usm.edu/jetde/vol12/iss1/2

[4] J. K. Hartshorne and L. T. Germine, "When does cognitive functioning peak? the asynchronous rise and fall of different cognitive abilities across the life span," vol. 26, no. 4, pp. 433–443. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0956797614567339

[5] S. Kang, E. Shokeen, V. L. Byrne, L. Norooz, E. Bonsignore, C. Williams-Pierce, and J. E. Froehlich, "ARMath: Augmenting everyday life with math learning," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, pp. 1–15. [Online]. Available: https://dl.acm.org/doi/10.1145/3313831.3376252

[6] E. Klopfer, *Augmented learning: research and design of mobile educational games*. MIT Press.

[7] E. Klopfer and K. Squire, "Environmental detectives—the development of an augmented reality platform for environmental simulations," vol. 56, no. 2, pp. 203–228. [Online]. Available: http://link.springer.com/10.1007/s11423-007-9037-6

[8] L. Y. Midak, I. V. Kravets, O. V. Kuzyshyn, J. D. Pahomov, V. M. Lutsyshyn, and A. D. Uchitel, "Augmented reality technology within studying natural subjects in primary school," p. 11.

[9] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information." vol. 63, no. 2, pp. 81–97. [Online]. Available: http://doi.apa.org/getdoi.cfm?doi=10.1037/h0043158

[10] C. Moro, Z. Štromberga, A. Raikos, and A. Stirling, "The effectiveness of virtual and augmented reality in health sciences and medical anatomy: VR

and AR in Health Sciences and Medical Anatomy," *Anatomical Sciences Education*, vol. 10, no. 6, pp. 549–559, Nov. 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/ase.1696

[11] H. Papagiannis, "Augmented reality (AR) joiners, a novel expanded cinematic form," in *2009 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media and Humanities*. IEEE, pp. 39–42. [Online]. Available: http://ieeexplore.ieee.org/document/5336728/

[12] H. L. Roediger and J. D. Karpicke, "Learning and memory," in *Encyclopedia of Social Measurement*. Elsevier, pp. 479–486. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/B0123693985005405

[13] A. S. Rose, H. Kim, H. Fuchs, and J. Frahm, "Development of augmented-reality applications in otolaryngology–head and neck surgery," vol. 129. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/lary.28098

[14] L. B. Rosenberg, "The use of virtual fixtures as perceptual overlays to enhance operator performance in remote environments." p. 53. [Online]. Available: https://www.researchgate.net/profile/Louis-Rosenberg/publication/235116787_The_Use_of_Virtual_Fixtures_as_Perceptual_Overlays_to_Enhance_Operator_Performance_in_Remote_Environments/links/5bd9a535a6fdcc3a8db385f5/The-Use-of-Virtual-Fixtures-as-Perceptual-Overlays-to-Enhance-Operator-Performance-in-Rem pdf

[15] L. Sherwood, *Human physiology: from cells to systems*, OCLC: 912318311. [Online]. Available: http://www.dawsonera.com/depp/reader/protected/external/AbstractView/S9781473732155

[16] P. Sommerauer and O. Mueller, "Augmented reality in informal learning environments: Investigating short-term and long-term effects." [Online]. Available: http://hdl.handle.net/10125/50063

[17] H.-K. Wu, S. W.-Y. Lee, H.-Y. Chang, and J.-C. Liang, "Current status, opportunities and challenges of augmented reality in education," vol. 62, pp. 41–49. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0360131512002527

[18] N. Xi, J. Chen, F. Gama, H. Korkeila, and J. Hamari, "The effect of operating in many realities on memory: An experiment on memory recognition in extended realities." [Online]. Available: http://hdl.handle.net/10125/71161

[19] H. Öğmen and M. H. Herzog, "A new conceptualization of human visual sensory-memory," vol. 7. [Online]. Available: http://journal.frontiersin.org/Article/10.3389/fpsyg.2016.00830/abstract

Appendix A

Supplemental Information

The following section will present and describe the apparatus used for the implementation.

A.0.0.1 Autodesk Maya 2022

Autodesk Maya is primarily used for this project to create the 3D models and UV-maps of the digits.

A.0.0.2 Unity

Unity is used to create the AR version of the game.

A.0.0.3 Visual Studio 2021

Visual Studio is used to script the behaviour of the AR game.

A.0.0.4 Adobe Photoshop

Adobe Photoshop will be used to create the non-virtual serial recall game. Creating posters with this software will be used to advertise the experiment.

A.0.0.5 AR devices

Equipment provided by Blekinge Institute of Technology IT helpdesk. The equipment provided was a Webcam (NAME).

A.0.0.6 Google Excel

Excel version on Google Drive will be used to store data collected from the experiment.

A.0.0.7 Vuforia

The platform used to create the AR game was Vuforia Engine(VERSION) which integrates image tracking. Vuforia is simple to integrate with Unity and keeps track of the recognition marker that would initiate the 3D models for the AR application.

A.0.0.8 Webcamera Logitech Tessar 2.0/3.7 carl Zeiss

This was the webcam that was used in our experiment, loaned by BTH's IT department.

A.0.0.9 Prism - GraphPad

This software was used to create statistic's, graphs and calculate the normality, homoscentricity, paired t-test and Wilcoxon signed rank test.

Appendix A

Consent Form

Consent Form

Consent form for participating in the Bachelor's degree project: Assessing the affect on short-term memory in students by comparing a serial recall Augmented Reality game and a card version.

The purpose of participating in this experiment is to help us to gather short-term memory data regarding the amount of digits recalled during two separate serial recall game versions, one being an AR game and the other a card game.

You may ask questions before and after the experimentation. The experimentation may take 20 minutes and offers no risk of mental or physical harm.

The experimentation will follow the next protocol:

1. Signing of Consent Form
2. Read the game rules
3. Play the first game version
4. Recovery time - 5 min
5. Play the second game version

The information gathered from you will be stored and encrypted until the completion of the bachelor degree project and will only be disclosed to the researchers and supervisor. The personal data collected will be: signature, name, location data and score.

By signing this consent form you hereby accept that your personal information will be managed within the limits of the study that is explained above. At any time you have the right to withdraw your consent without giving a reason and contacting any of the studies contact personnel according to the information paper. Your personal information and data will thereby be deleted and no longer be accessible or stored.

☐ You consent to taking part in this experiment.

☐ You consent to have your information stored until completion of the bachelor degree project.

☐ By crossing the mark you acknowledge that you are physically healthy and have not experienced any covid symptoms from the last two weeks.

....

Signature ***City and Date***

.....

Name clarification

You always have the opportunity to get information about what has been registered about you or if you have views on the process or the information collected by contacting one of the persons below:

Student name: Linus Dorell
Telephone number: 0705408878
Email address: lido18@student.bth.se

Student name: Oskar Nyman
Telephone number: 0702486122
Email address: osny19@student.bth.se

Supervisor name: Diego Navarro
Email address: dna@bth.se

Faculty of Computing, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden