Notetaking in Minecraft and the Educational Use of Video Games: An Experimental Study

Szilard A. Reynisson

Bacheloroppgave i psykologi

Veileder: Francesco Margoni

Stavanger, 25.04.2023

Det samfunnsvitenskapelige fakultet

Institutt for sosialfag

Word count: 7393

Acknowledgments

I wish to thank Martine and Basse, for their support through the entirety of this project. I also want to thank my family, and all those who encouraged me to feed my curiosity and pursue this topic. A special thanks goes to my guidance counsellor, Francesco, for giving the green light on this project, and for always providing useful advice and guidance, whether I agreed with it or not. I wish to thank the teachers at Sola, Sandnes and Dalane VGS for making this possible, and all the students who took the time to participate. And finally, a word of appreciation for The University of Stavanger, for making any of this possible.

0. Abstract

By no means is digital game-based learning a recent idea, yet empirical evidence surrounding it is hazy. This paper aims to attain clarity and understanding of the topic and contribute to the literature. This contribution came in the form of an experiment comparing Norwegian high school students' knowledge test scores, after studying a text on memory using three different techniques. One of these required the use of Minecraft Education, the intervention. The remaining two being controls. Self-reports regarding task attitudes were collected. No significant effect was found between the use of Minecraft Education and test scores, but significant effects were found for task attitudes. Besides motivation being a strong predictor of test score, these effects imply that study technique could have indirectly influenced scores as well, providing interesting possibilities for future research.

Dataspill-basert læring er ingen ny idé. Likevel er forskning rundt det uklar. Denne studien forsøker å forstå feltet bedre, og bidra til det. Det ble gjort ved å sammenligne resultater på en kunnskapstest, utført av elever ved norske videregående skoler. De ble delt i tre grupper, hver med sin studieteknikk å lære innholdet fra en tekst med. En av disse brukte Minecraft Education, intervensjon gruppen. De to andre var kontrollgrupper. De ble også spurt om holdninger til oppgaven. Ingen signifikante funn mellom studieteknikk og testresultater ble funnet, men flere signifikante forhold ble funnet mellom holdninger. Foruten om en sterk korrelasjon mellom motivasjon og testresultater, antyder disse forholdene at studieteknikk kan ha indirekte påvirket testresultater. Dette presenterer spennende muligheter for fremtidig forskning.

Contents

1. Introduction

1.1 Recent Research

Video games are mostly thought of as a form of entertainment, but a body of literature has begun to emerge exploring how this form of media could be utilized in more constructive manners (Granic et al., 2014). This is welcome news, considering how much of the existing literature is focused on the negatives of these games. Four major domains within which this recent research has been focused are the cognitive, motivational, emotional, and social domain. (Granic et al., 2014,). This paper mainly belongs to the cognitive domain, as I attempt to explore how video games might find their place in the classroom.

Within the cognitive domain a popular paradigm in modern research is exploring how video game playing can lead to improved cognitive abilities, such as executive functions, information processing, visual processing skills etc. (Powers et al., 2013,) This is often done by either comparing people who play video games on their free time with those who do not, or by intervention studies. Results indicate that there is indeed a link between video game playing and increased cognitive abilities (for meta-analyses see Powers et al., 2013; Bediou et al., 2017). However, meta-analysis results are likely affected by publication bias, with effect sizes estimated to be artificially inflated by up to 30%. There is also the argument that while video games seem to have the ability to improve cognitive abilities, this improvement is task specific and hardly transferrable to other applications (Sala et al., 2018).

As for educational applications of video games, more specifically entertainment video games, research is sparse. This is illustrated by a meta-analysis finding a total of 49 intervention studies conducted between 2005 and 2019, concluding that there is too little data in general, but especially quantitative data on this topic (Martinez et al., 2022). The existing literature is

however in the favor of these type of interventions, and there also exists a somewhat larger body of literature concerning the topic of serious games. These are video games solely dedicated to the purpose of learning or training, without providing entertainment value (Gao et al., 2020). This topic is explored in a systematic review by Ravyse et al. (2017), which analyzed a total of 397 articles. Their findings suggest that while serious games are unlikely to encourage their players to read up on the intricacies of agriculture after a session on Farming Simulator, their knowledge of the content material does increase. This is especially likely to occur when players enter a flow state (Csikszentmihalyi, 2000), which these games are likely to elicit to various degrees, moderated by certain factors. One of these factors is how entertained players are by the game, despite the design of serious games. Another systematic review further suggests that digital game-based learning has potential within education, specifically within content understanding; This echoes the sentiment of other papers on this topic, in that there is a need for more research. (Hussein et al., 2019)

1.2 Historical Perspectives

In an interesting alternative to the perspectives outlined above, Simon Egenfeldt-Nielsen had this to say in his overview of the educational use of video games, written in 2006 “More than once we have heard that research on video games is an emerging field in which there has been no prior research, even though this is clearly not the case. Unfortunately, amnesia shackles too many researchers.” (Nielsen, 2006). This overview explores more than 300 references on the topic, which provide behavioral, cognitive, constructivist and socio-cultural perspectives on the subject.

And indeed, heavily cited articles from the same era by Squire et al. raise the same point. Educators are ignoring the potential benefits of video games, or they are using educational video game content that is lacking in empirical evidence. The bottom line is that both the

commercial and military sector, digital game-based learning is receiving significant attention, and the educational sector must catch up. (Squire, 2003; Shaffer et al., 2005).

Considering this, one would expect that by today, there would be more conclusive evidence on how, which, when and why video games have a place in the classroom. Surprisingly, this does not seem to be the case. Are researchers still shackled by amnesia, as Nielsen put it, or is there some other explanation for this? I believe he provides at least one part of the answer to this in the conclusion of his overview, stating that while all the perspectives he explored proved valuable insight, there are theoretical inconsistencies between them. Additionally, he points to a lack of a clear-cut approach to researching the topic (Nielsen, 2006). Perhaps it is not so much a lack of research, or interest thereof that hinders advances in digital game-based learning. Rather, it might partly be due to a lack of a coordinated effort or clearly defined research paradigm or culture.

1.3 Research Goals

With this paper my aim was to explore the viability of one possible approach to researching digital game-based learning. Specifically, I wanted to explore whether taking notes using an entertainment video game, in this case Minecraft Education (Mojang, 2023), will lead to a higher degree of recall, as measured by a knowledge test, among high school students than traditional paper-based note taking or simply reading and rereading a text will. Of available edutainment games, Minecraft Education was deemed a good option for multiple reasons. Firstly, everyone with an email address issued associated with the public educational institution in Norway can access the game freely. There is also a relatively good amount of existing literature on the game (Nebel et al., 2015), which this study aims to build upon. Furthermore, Minecraft Education and even the non-educational edition of Minecraft have a near endless capacity for creative expression.

Minecraft is a three-dimensional building game, that can be both a single and multi-player experience. In its' most basic form, the gameplay is centered around placing or removing a variety of blocks, three dimensional pixels, within this space. This can easily be grasped by kids as young as three years of age (Mavoa et al., 2017). However, it can also get extremely complex when players begin taking use of the game's version of electrical engineering, a system called Redstone. Examples of this are Scicraft, an exclusive server played by professors and students of fields like computer science, engineering, and machine learning. Their goal? Building machines that automate all aspects of the game (Mumbojumbo, 2020). Another perhaps even more impressive example is an ongoing trend where players compete to build the most complex in-game computer using the Redstone system. One such computer was able to run a simplified version of Minecraft, within Minecraft (Sammyuri, 2022). Minecraft Education even more tools, like programming and real-life chemistry, making it a truly versatile tool in the classroom.

Only a couple steps back from in-game engineering, this study asked participant to "take notes" from a text by building simple structures within a single-player Minecraft Education world. These structures were meant to symbolize various pieces of information from the text, to aid recall when completing the upcoming knowledge test. A study in similar spirit was conducted among elementary school children from a high socio-economic area in Taiwan. Children were placed in one of two groups, one where they were instructed to read a website about firemen, and another where they played a firefighting-video game. The information presented about the subject was the same in both conditions, and the children from the video game condition achieved significantly higher scores on a subsequent knowledge test (Chuang & Cheng, 2009). In another study, military personnel were instructed to study and train for a test on a particular subject matter. The first group was handed a text, to read as they wished. The second group received a text and an accompanying test with which to train, while the

third group were given access to a video game about the subject matter. During subsequent retention testing, not only did the participants from the gaming group score higher, but they also rated their task more enjoyable and effective than participants in the other groups did (Ricci et al., 1996).

Digital game-based learning is evidently not a new idea, yet the research is so inconclusive. After all, how on earth would playing games in school be a good idea? As a matter of fact, there are several ways this could lead to a net increase in learning. We know that video game playing and cognitive abilities correlate, but since these cognitive ability increases may be domain specific, it is unsure whether they would lead to students learning more. However, video games possess a quality few other types of media do: an interactive multisensory experience. In addition to engaging both hearing and vision, like video-based media does, video games require the consumer to engage with them to take an active role. Players must interact, manipulate, and contemplate the elements of the game. In theory, this forces them to process material on a deeper level (Holt et al., 2019), while they also stay engaged, provided they reach the flow state discussed earlier.

The argument for multisensory learning is that in early life learning is highly multisensory, and while this reliance on multiple senses for learning is reduced in adulthood, it does not disappear (Shams & Seitz, 2008). Seitz et al. (2006) have demonstrated that when assigning participants to either a visual, or audiovisual group, and training them for a coherent motion and discrimination task, the audiovisual group had significantly higher rates of learning when compared over multiple sessions than the visual group. Multisensory learning is among other things a foundation of the Newton Project, an EU backed initiative to research and build up technology enhanced learning (TEL) (The Newton Project, 2023).

In addition to the multisensory argument, video games are known to produce emotions (Hemenover & Bowman, 2018), which can influence learning, and by extension memory, in

several ways (Tyng et al., 2017). Emotional events are for instance remembered more clearly, precisely and for longer than emotionally neutral events. Of course, emotions can be both positive and negative. This is reflected in the Circumplex Model of affect (Russel, 1980) which holds that all emotions are an interplay of two spectrums, affect and valence. Studying how emotion affects learning through this model, has suggested that negative emotional arousal leads to decreased associative learning. However, positively arousing emotions have recently been shown to do the opposite (Madan & Kensinger, 2019). This suggests yet another mechanism that digital game-based learning might capitalize on to aid students in recalling their curriculum.

To summarize, the educational use of video games is a relatively little explored, inconclusive, and scattered field of research. The bulk of the existing research is focused on serious games, which are games specifically designed for education and training. I made the argument that entertainment video games may have qualities that make them suitable for teaching, based on evidence from the research on serious games, digital game-based learning, multisensory learning and on the moderating effect of emotions on memory. This leads to my hypothesis. I predict that high school students (typically aged 16-19) who take notes from a text using Minecraft Education will subsequently attain higher scores on a knowledge test than students taking notes by hand, or students who only read the text (H_1). My null hypothesis (H_0) predicts no effect at all of study condition on the test score. In addition to the knowledge test, all students would be asked control questions about their motivation, enjoyment, excitement, and preference for the task. This was to control for moderating effects on test score and to exclude unserious responses.

2. Methods

2.1 Literature Review

For the literature review done in advance of this experiment, Google Scholar (Google, 2023) was used as the primary search engine. Keywords such as “games and learning, digital learning, video game learning, video games and cognition, educational video games, Minecraft education, video games and cognition, cognitive effects of gaming, video games effects on memory, video games and recall” and other similar terms were used. In addition to google scholar, the websites Paper Digest (Paper Digest, 2023) and Elicit (Elicit, 2023) were used. These are tools built for researchers meant to aid in performing literature reviews, by searching, analyzing, and summarizing literature from multiple scientific databases on any given topic. The final tool used in this process was Litmaps (Litmaps, 2023), a similar tool that creates a “web” when fed any research paper, visualizing relations with other scientific publications.

2.2 Participants

The population of interest for this study were Norwegian high school students from Rogaland county. Participants were recruited from three separate schools, all of which volunteered students from their first-year psychology classes, making this a convenience sample. At this stage there were 76 willing participants, but due to reasons outlined below the final amount was N = 35, aged 17-24 (M = 17.65, SD = 1.53) being 77% (n = 27) female. The reading group was somewhat larger, (n = 20), than the gaming group, (n = 15).

There were three levels of the independent variable (study technique) in this study, and each of the three schools were assigned to one of these. The first school comprised the reading condition which acted as the primary control group, measuring performance in the absence of active engagement with the reading material. The second school made up the gaming condition, the intervention group. The third and final school was assigned to the note-taking

condition. The purpose of this second control was to rule out active engagement with the reading material as a confounding factor on subsequent knowledge test scores. By this design, internal reliability was deemed high enough to be confident in any potential effects that would be found.

2.3 Materials

2.3.1 Reading Material and Knowledge Test

One condition for being allowed to recruit this group of participants was to make the materials of the experiment relevant to the student's curriculum. All schools followed the same plan, and the next topic they would be learning about in class was memory. Therefore, I authored a text about memory, based on chapter 8 on memory from Psychology: The Study of Mind and Behavior (Holt et al., 2019), a university level psychology introduction book. The text was approximately 1000 words over two pages and was administered to the participants in paper format. Every single piece of unique information in the text was assigned a number. This was used when grading participant responses. Each piece of information recalled awarded students one point, out of a maximum score of 79 points. This open-ended format made grading each response accurately more of a challenge. However, it provided the benefit of reducing ceiling effects, and chance-guesses associated with multiple-choice questionnaires. This was deemed as a reasonable trade-off.

The inspiration for this knowledge test was primarily based on two studies. The first of which is an experiment examining whether reading the same text on screen or paper has any effect on reading speed, and text comprehension. One of the tasks in the subsequent text comprehension test was an open-ended question, asking participants to list any information they could remember from the text. This response would then be scored based on a predefined list of statements, where each correct statement listed would grant students one

point (Ronconi et al., 2022). The second study examined how different methods of notetaking would affect recall of a lecture. In this study too, recall was measured using a variety of tasks. One of these tasks was a free recall task, where participants were given 10 minutes to recall as much as they could remember from the lecture. Their responses were then scored by two independent raters, assigning points for each "idea unit" participants could recall (Bui et al., 2013).

To ensure the reading material was at an appropriate level, a readability index was calculated. In Norway, this is commonly done using the "lesbarhetsindeks", (liks) a formula created by Carl-Hugo Björnsson during the 60s (Store Norske Leksikon, 2020). It uses the length of words and sentences to calculate a score, commonly in the range of 20 to 60 points, with 40 being considered a "normal reading level". The liks of the reading material was calculated to be a modest 36. In comparison, the mean liks for ten random texts sourced from NDLA (NDLA, 2023) came out at 43.2. Calculation was done using an online liks calculator (Nordtømme, 2023).

2.3.2 Distraction Task and Response Forms

A distraction task was also prepared, in the form of a quiz using the online software Kahoot (Kahoot, 2023). The quiz was comprised of 20 questions, half of which were psychology-related, the other half pop-culture related. The type of questions consisted of multiple-choice questions, with 1-2 correct answers out of 4-5 alternatives. There were also true/false questions, rank-based questions (e.g., arrange these states from Eriksson's psychosocial development theory in order of occurrence) and one question where students had to point out the location of the occipital lobe on a brain-structure diagram. One question was unique for each school, asking the participants to select the name of their headmaster out of four alternatives. Kahoot puts participants in direct competition with each other and showcases

scores and rankings after each question. These scores are based on points acquired both by answering correctly, and by quickly submitting the answer. The stress generated by this, combined with a variation of questions and tasks, requiring substantial levels of attention and cognitive effort, all while being fun and engaging, are the reasons for selecting Kahoot as the distraction task.

Participants were handed an electronic form in two parts within which they gave their informed consent, as well as their response. This form was downloaded by the participants from their electronic school portal in advance of the experiment. They were asked to fill out the form with their participant number, gender (male/female/other) and their age in years. In part one of the form, students wrote down as much information as they could remember from the reading material. In part two they were asked to answer how strongly the agreed with four questions on a Likert-scale, with intensity ranging between 1: “I completely disagree” to 5: “I completely agree”. These questions were: “I was motivated to perform well on this task”, “I found this task to be exciting”, “I found this task to be entertaining”, and “I prefer this method when I study on my own”.

Materials unique to the experimental group consisted of a world in Minecraft Education for showcasing purposes, and a text on Electromagnetic Radiation sourced from the educational database NDLA (NDLA, 2023). All materials can be seen in the appendix.

All materials were in Norwegian.

2.4 Equipment

A complete list of equipment used includes two laptops (one for connecting to the projector, one as a backup/to lend participants), a projector/smartboard, a full set of reading material in paper form, a full set of paper and pencils for the notetaking condition, 80 unique pieces of

paper with participant numbers on them, an opaque bag, a scissor and two USB memory sticks.

2.5 Procedure

2.5.1 Video Game Course

The procedure was the same for each school, but the experimental group also received a 45-minute course on using Minecraft Education, two days prior to the experiment. This course taught participants the basics of using the software, and the specific study technique they were to use on the day of the experiment. To achieve this a text about electromagnetic radiation was used as an example of how to take notes in Minecraft Education. Using as many different blocks as possible, I constructed multiple structures in the example world, each of which symbolized a unique set of information from the text. Students were taken on a brief tour of this world, following along by watching on a smartboard. They were explained how these structures related to the text on electromagnetic radiation, to encourage, inspire and give reference as to how they could create their own study worlds. This tour lasted roughly 20 minutes and was followed by 20 minutes where participants were given an opportunity to try out the method themselves. They were instructed to choose any text they wanted from NDLA and were free to ask me or my research assistant for help or guidance. The final five minutes were spent showing the class how to export these worlds and submit them anonymously on the day of the experiment.

2.5.2 Experiment

The first ten minutes were spent briefly introducing myself and the experiment, randomly assigning participant numbers by drawing a number from a bag, handing out the materials and making sure everything technical was in order. After this came the main part of the

experiment, the study task. Each group was given 25 minutes to study and memorize the contents of the text to the best of their ability, using the study technique they had been assigned to. Here the reading group simply read the text over and over, using whatever reading technique they preferred. The note-taking condition was given an additional piece of paper and a pen, and told to take notes, draw, make mind-maps or whatever else they should prefer. The gaming condition studied using the technique they had previously been taught. After the 25 minutes had passed, the reading materials and notes/Minecraft Education Mode study worlds were collected so participants no longer had access to them.

Following this came the distraction task, to which an approximate 15 minutes were assigned, due to the somewhat variable length of each quiz. Students received a PIN code with which they could connect to the quiz. Kahoot shows how many devices are connected to the quiz session, and by comparing this to the number of students in the class it was ensured that each participant was connected. As an additional safety measure both a teacher and my research assistant did laps around the classroom, ensuring that participants did not cheat by opening web pages or any other means during the experiment. Doing so would cause them to be excluded from the study, of which they were informed.

Following this, participants were given 25 minutes to write down as much as they could possibly recall from the study material. They were encouraged to write down absolutely everything that came to mind, even single words. During this part participants could ask the researchers practical questions about the process itself but received no further assistance. They were not allowed to communicate with each other either.

Finally, five minutes were allocated at the end to collect the participant forms. To ensure anonymity, these were collected by a research assistant, who distributed USB memory sticks to the students to which they could transfer their response forms. The forms were to be marked with their participant number, and nothing else. We quickly realized that this method

of response collection was flawed, as most participants were using computers without USB ports. To make up for this, participants were instructed to set the file name of their response sheets to nothing but their participant numbers and send them as an email attachment to their teachers. They would then forward the files to me, ensuring participant anonymity.

2.6 Statistical Analysis

2.6.1 Attrition and Exclusion

Out of the initial 76 participants, only 35 were included in the final dataset. This was in part because some responses from the reading and gaming condition were submitted to the wrong email, corrupted or just simply never submitted. A majority of these were recovered, but some were lost for good. Others were excluded on suspicion of not taking the experiment seriously, based on abnormally low knowledge test scores. The basis for this exclusion criteria was that the average human can maintain seven, give or take two, meaningful units of information in their working-memory (Holt et al., 2019). With that in mind, a score below five points was deemed as a reasonable cutoff. Also, several participants never showed up to the experiment. This was due to sickness, and a school trip some participants from the gaming condition were on. The primary cause of attrition however was that the entirety of the note-taking condition had to be discarded. The experiment was postponed once, and by the time it was completed, this class had already had a lecture on the exact contents of the reading material.

2.6.2 Variables

Participant data was collected for a total of 11 variables. Three of these were specific to the intervention group and were excluded from final analysis as they were deemed redundant. Out of the eight that remained, the IV was study technique, a nominal variable. The primary dependent variable (DV) was recall, operationalized as knowledge test score, a ratio variable.

The attitude questions about participant motivation, excitement, task enjoyment and preference for assigned study technique were all interval variables. These served as additional DV's, and control variables. Participant age and gender was also collected, being an interval and a nominal variable. These too acted as control variables. All variables within this study were discrete. See appendix for further definitions and values.

2.6.3 Analyses

All statistical analyses were performed using SPSS version 28.0.1.1. As all data were discrete, no test of normality was performed. Due to this, the Mann Whitney U test was selected, as it is a non-parametric test suitable for comparing independent samples. This is because instead of operating with group means like a T-test would, it uses a ranking system based on group median scores. The test was performed twice, first with IV level then with gender as the grouping variable. Following this I performed a bivariate correlations analysis, followed by an exploratory factor analysis to control that the attitude questions covaried, before further subjecting them to a reliability test. This resulted in a new variable called attitude, a combination of participant excitement, enjoyment, and motivation. Following this the Mann Whitney U test and the correlational analyses were performed a final time, this time including the new variable. Mann Whitney U test effect sizes were calculated manually as SPSS lacks this function. These effect sizes were calculated using Wilcoxon's r. Descriptive statistics were also calculated in SPSS, including group-by-group median scores to use in the Wilcoxon's r calculations.

3. Results

3.1 Descriptives

Table 1:

Descriptive Statistics by Group

		Recall	Motivation	Excitement	Enjoyment	Preference	Age	Gender	Attitude
Reading	M	17.80	3.55	3.50	3.35	3.00	18.11	1.75	3.47
	N	20	20	20	20	20	19	20	20
	SD	10.36	.83	1.28	1.18	1.08	1.94	.44	.93
	Mdn	16.50	4.00	4.00	4.00	3.00	17.00	2.00	3.67
Gaming	M	17.93	3.46	3.92	4.08	1.85	17.07	1.80	3.82
	N	15	13	13	13	13	15	15	13
	SD	8.44	.97	1.04	.95	.80	.26	.41	.81
	Mdn	16.00	3.00	4.00	4.00	2.00	17.00	2.00	3.67
Total	M	17.86	3.52	3.67	3.64	2.55	17.65	1.77	3.61
	N	35	33	33	33	33	34	35	33
	SD	9.45	.87	1.19	1.14	1.12	1.54	.43	.89
	Mdn	16.00	4.00	4.00	4.00	3.00	17.00	2.00	3.67

As seen in Table 1, initial descriptive analysis revealed that the gaming condition scored half a point lower (Mdn = 16.00, SD = 8.44) than the reading condition (Mdn = 16.50, SD = 10.36) on the knowledge test. The reading group also scored a point higher on motivation (Mdn = 4.00, SD = .83) and preference for their assigned study technique (Mdn = 3.00, SD = 1.08) than did the gaming group (Mdn = 3.00, SD = .97) (Mdn = 2.00, SD = .80). Median scores for the remaining variables were identical between the two groups, with minimal variance in deviation (SD = .19 – SD = .28). The only exception to this was age, where there was a SD = 1.68 variation between the gaming (SD = .26) and the reading group (SD = 1.94).

3.2 Factor Analysis and Reliability Test

Table 2 shows the results from the exploratory factor analysis, which yielded two factors. Factor 1, as seen in table 2 explained 45.32% of variance and loaded between .937 and .458. Factor 1 consisted of task enjoyment, excitement, and motivation. Factor 2 explained

Table 2:

Pattern Matrix and Eigenvalues From EFA

Item	Factor	
	1	2
Enjoyment	.937	-.313
Excitement	.775	.362
Recall	.044	.726
Motivation	.458	.562
Preference	-.036	.161
Initial Eigenvalues		
Total	2.27	1.22
% of Variance	45.32	24.45
Cumulative %	45.32	69.78

24.45% of variance, loading between .726 and -.313. Several items were cross loaded. Due to the low eigenvalue and high cross loading, factor 2 was discarded. The reliability test for the three items in factor 1 yielded a good Cronbach's Alpha ($\alpha = .79$). Removing items would only have decreased α (see appendix). These items were combined, creating a new variable.

3.3 Mann Whitney U Test

The Mann Whitney U test was performed using both the IV study technique and gender (see appendix) as a grouping variable, examining group variations in recall/knowledge test scores

and attitudes. Only when grouping by condition were significant results observed, but none between the knowledge test scores (U = 160.00, Z = -.69, p = .49, r = .16) of the gaming (Mdn = 16.00, n = 15) and reading (Mdn = 16.50, n = 20) condition. As illustrated in Table 3, there was a small, but significant variation between the gaming (Mdn = 4.00, n = 15) and the reading (Mdn = 4.00, n = 20) group when looking at task enjoyment (U = 106.50, Z = -2.07, p = .04, r = .16). The same variation was seen between the gaming (Mdn = 2.00, n = 15) and the reading (Mdn = 3.00, n = 20) group when examining their preference for their assigned study technique (U = 73.50, Z = -3.08, p = ≤ .01, r = .16). Further confirming what the descriptive analysis suggested, a variation in age was observed between the gaming (Mdn = 17.00, n = 15) and the reading condition (Mdn = 17.00, n = 20) group (U = 128.50, Z = -2.10, p = .04, r = .16).

Table 3:

Mann Whitney U Test With Estimated Effect Sizes

	Recall	Preference	Attitude	Age	Motivation	Excitement	Enjoyment	Gender
U	160	73.5	138	128.5	161.5	138.5	106.5	180
Z	-.686	-3.076	-1.038	-2.097	-.344	-1.051	-2.072	-.262
p	.493	.002	.299	.036	.731	.293	.038	.793
r		.16		.16			.16	

[a]r = Wilcoxon's r

[b]grouping variable = IV level

3.4 Correlations

As can be seen in Table 4, a significant, positive correlation of medium strength was found between the IV and task enjoyment, r(33) = .34, p = .04. This indicates that students were more likely to enjoy themselves when assigned to the gaming condition. The IV also had a significant, negative correlation of medium strength with preference for the assigned study technique, r(33) = -.50, p = ≤ .01, suggesting a tendency to prefer reading over gaming as a study technique. The biggest predicter of knowledge test score was motivation, correlating positively with medium strength, r(33) = .47, p = ≤ .01. Task excitement also had a medium strong positive correlation with test scores, r(33) = .39, p = .02. This shows that students who were excited to be a part of the study and were motivated to do well on the task tended to

Table 4:

Complete List of Correlations Between Variables

	Variable		1.	2	3	4	5	6	7	8	9
1	Condition	r									
		p									
2	Recall	r	.065								
		p	.693								
3	Motivation	r	-.051	.473**							
		p	.760	.003							
4	Excitement	r	.203	.389*	.662**						

		p	.222	.017	≤ .001					
5	Enjoyment	r	.336*	-.089	.346*	.641**				
		p	.039	.602	.033	≤ .001				
6	Preference	r	-.497**	.092	.232	.036	-.009			
		p	.001	.590	.162	.832	.957			
7	Age	r	-.307	-.142	.040	.129	.071	.179		
		p	.057	.396	.815	.447	.678	.288		
8	Gender	r	.043	.064	-.267	-.093	-.039	-.204	.168	
		p	.797	.702	.109	.584	.818	.226	.314	
9	Attitude	r	.206	.303	.778**	.928**	.802**	.092	.009	-.150
		p	.215	.068	≤ .001	≤ .001	≤ .001	.581	.559	.376

*p ≤ .05. **p ≤ .01.

score higher than their less motivated and excited peers. Unsurprisingly, motivation and excitement correlated positively, r(33) = .66, p = ≤ .01, demonstrating a fairly large effect size. This trend indicates that excited students were likely to be motivated, and vice versa. Motivation also positively correlated with task enjoyment at medium strength, r(33) = .35, p = .03, suggesting that more motivated students also enjoyed themselves more. Finally, also task excitement and task enjoyment shared a strong, positive correlation, r(33) = .64, p = ≤ .01., showing an increased likeliness to enjoy the task when the participant was excited. Also worth noting is that besides the expected, but only marginally significant correlation with study technique, r(34) = -.31, p = .06, age had no effect with any variable. See also Figures 1 and 2 for attitude effects.

Figure 1:

Scatter Plot With Fit Line Showcasing Condition (y) and Participant Attitude (x) Interactions

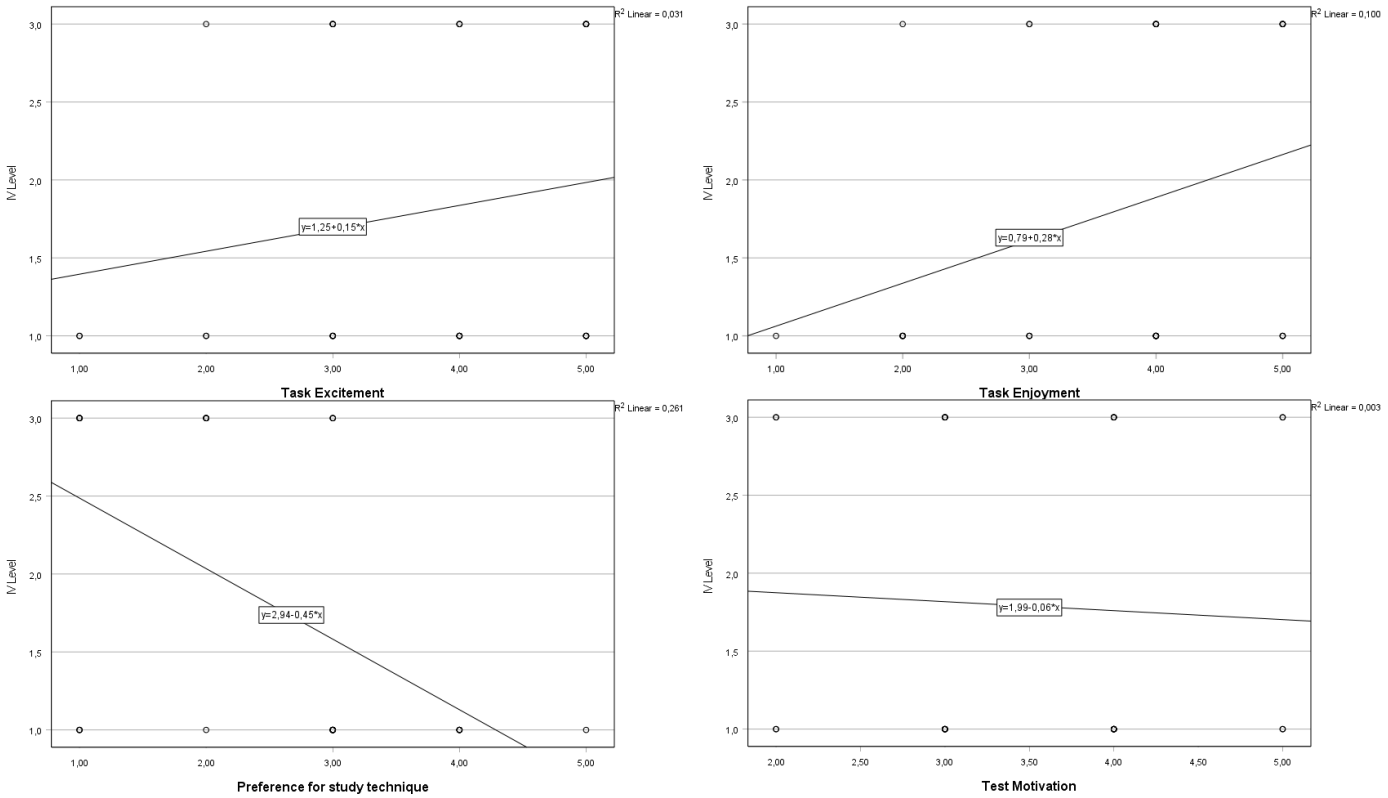
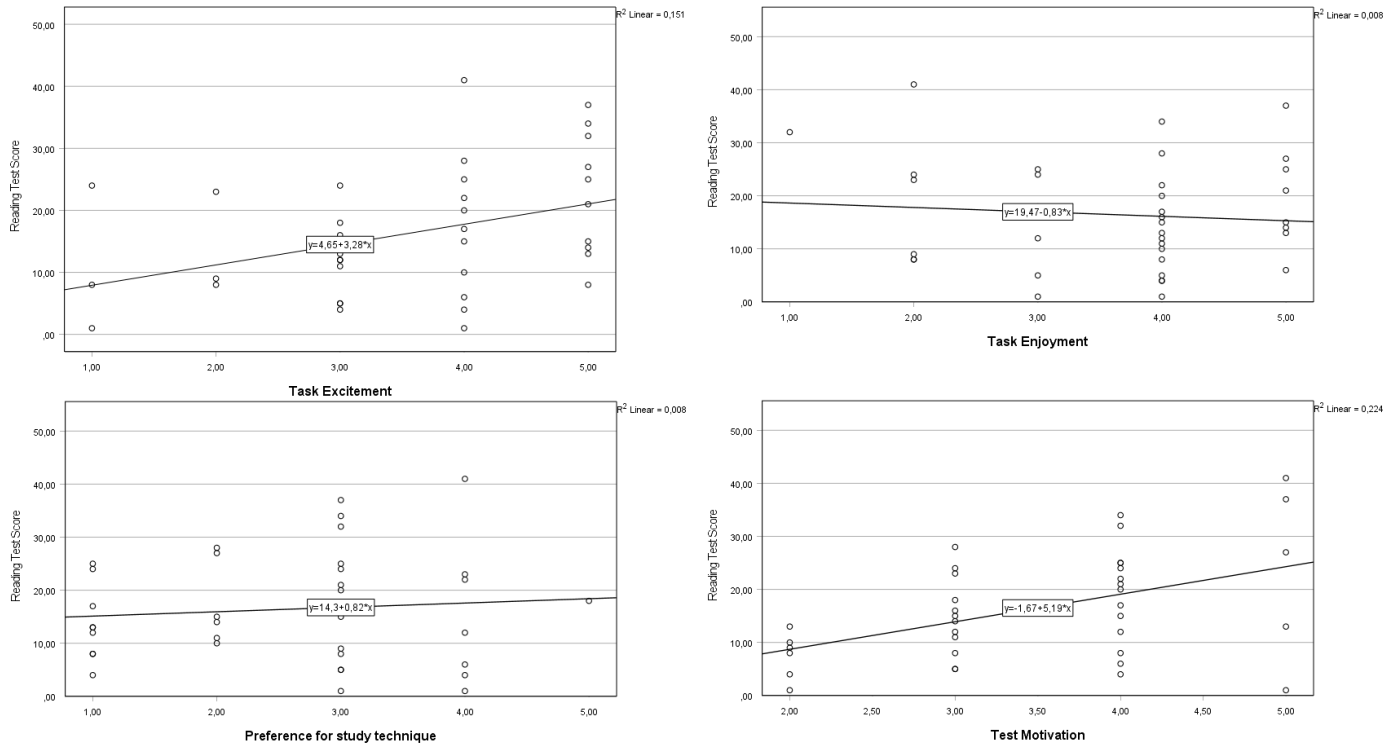


Figure 2:

Scatter Plot With Fit Line Showcasing Test Scores (y) and Participant Attitude (x) Interactions



4. Discussion

4.1 Hypothesis and Research Goals

The primary goal of this study was to examine if taking studying a text aided by Minecraft Education would lead to higher recall than using more common study techniques. H_1 predicted that the students taking notes within Minecraft would achieve higher scores than their peers in the notetaking or reading condition on a subsequent knowledge test. And while the notetaking group was unfortunately discarded, the gaming group did in fact have a slightly higher mean knowledge test score than their reading group counterparts. This however was the only "direct link" observed between the IV and test score. A significant relationship was neither found during the Mann Whitney U test, or the correlational analysis. This falsifies H_1 and supports H_0, which predicted no effect between the IV and test scores.

However, the general goal of the study was to explore the educational potential of digital game-based learning. And indeed, several interesting and significant effects were observed which might give valuable insights to this topic. To begin with, the positive correlation between study condition and task enjoyment shows that student who played the video game were more likely to enjoy themselves. This relationship can also be seen in the Mann Whitney U test. Furthermore, task enjoyment correlated positively with both excitement, and motivation for the task. Besides correlating with each other, both variables correlated positively with test scores. Despite this, students tended to favor the reading condition more strongly as their preferred study method.

These findings suggests that despite not having a direct effect on test scores, the IV may still have influenced recall, in its relationship with task enjoyment. Task enjoyment may in turn have acted as a moderating variable, through its relationship with task motivation and excitement. The implications of this are that playing video games is enjoyable, and students who enjoy themselves are more likely to learn. Exploring these implications is however beyond the scope of this study.

4.2 Other Findings

Implications aside, several significant correlations were found (Table 1). Out of these excitement and motivation had the largest effect size, with excitement and enjoyment coming in at a close second, both being strong correlations. Moderate correlations were also found between test score and motivation, test score and excitement, motivation, and enjoyment, as well as the study condition and enjoyment and study condition and preference thereof. Without speculation, this shows that self-report of motivation is related to being excited, enjoying a given task and achieving higher scores on a knowledge test. Additionally, related to both test scores and enjoying an assigned task is being excited about that task. Lastly,

correlating negatively with preference for assigned study technique but positively with enjoying your task is the study condition.

What can be taken away from this is either that motivated students find more excitement and fun in doing tasks, or that students who enjoy and get more excited about their tasks also seem to be more motivated to perform well. Either way, these qualities are also linked to higher knowledge test performance. Whether they are the cause of this, or because students who expect to do well tend to be more motivated and excited, is unclear. What is clear however, is that students do not prefer Minecraft Education as a study technique, but they do enjoy using it. Perhaps because it is an engaging way of learning, perhaps because they got away from the predictable monotony of their school day, perhaps for any number of other reasons. The only sure thing is that these students reported having more fun. In any case, these results give some valuable insight to which factors might play into student motivation and test scores.

4.3 Strengths and Limitations

4.3.1 Comprehensive Design

By design, this was a randomized, controlled trial. By contacting every high school in the county, the entire population of interest was given an opportunity to take part in the study. Had more schools accepted, sampling would have been done by multistage sampling. In total three schools volunteered each of their first-year psychology classes to partake, making this selection of candidates a convenience sample, and biased. Regardless, each class was randomly assigned to the intervention, or one of two control conditions. During the experiments it was made sure that participants received the exact same treatment as well, stretching as far as standardizing everything except the physical location of each classroom. Also, by opting for a single-session design with a distraction task, so risk of some participants

discussing, maintaining, or researching the test material while others did not was eliminated. This way, despite the biased sample, the study remained both randomized and controlled.

It had to be ensured that participants were measured fairly. This was achieved by taking inspiration from existing knowledge tests, which measure performance on similar tasks and ensuring that the prepared reading material was of an appropriate level of complexity. Due to the length of the text, and design of the test, ceiling effects and chance responses were ruled out. Additionally, participants had the opportunity to ask for clarification at any time, making sure everyone had equal opportunity to perform well.

4.3.2 Room for Improvement

Despite these measures, obstacles were encountered. During data collection, some participants appeared noticeably less interested in the experiment than the rest. The 5-point exclusion criteria and self-report attitude questions were meant to filter out their responses, but it cannot be certain that all participants applied themselves equally. The reasons behind this behavior are unknown, but some explanations may be that these participants felt forced to be there, they did not perceive the research situation as “real” or “serious” enough, or the text could have appeared overwhelming. Perhaps these participants always act this way. These issues could potentially be mitigated by having a sample large enough to negate outliers, inviting participants to a dedicated testing facility to give the experience a more authoritative appearance, rewarding participation with a minor monetary incentive or shortening the reading material. The highest score any student achieved was 41 of 79 points, or meaning they recalled 52% of the information in the text (see appendix). This suggests the material was excessively long.

A simpler, more practical issue arose during the collection of finished response sheets. These were meant to be collected on USB drives to ensure anonymity, and not be dependent on an

internet connection. Unexpectedly, few participants had experience with this method of data transfer, and most were missing USB drives on their laptops. This caused some confusion, and other solutions had to be improvised. During future research, an online survey tool is likely to cause less issues.

I received help from a research assistant, but only in data collection. Due to this, I was aware of participant's study condition while rating their response sheets. This is a potential source of researcher bias. Going forward, this might be improved upon by employing one or more dedicated raters. Alternatively, collecting responses to the knowledge test separately of other data.

Lastly, as this study was conducted for a bachelor's thesis, there was a limited amount of time to prepare the experiment. Besides putting constraint on how carefully the reading material, the knowledge test or any other material could be crafted, this presented a significant challenge regarding coordination. With more time to discuss and plan together with the participating schools, several conflicts could have been avoided. The school trip could have been planned around, and the issues surrounding the notetaking school could potentially have been resolved. For instance, preventative measures like having substitute teachers available to avoid postponing could have been employed. Alternatively, reading material could be entirely unrelated to the curriculum. Another potential limitation caused by time constraint is the length of training received by the intervention group. The 20 minutes participants had to practice the study technique were not enough. It is my belief that by giving these participants more time to master the gaming strategy, the likelihood of significant variations of recall would have been higher between the groups. With more time to plan, these issues will all be negatable during future replications.

4.4 Future Research

In the above section I outlined how this study might be improved upon in future replications, but does it raise any new questions? For one, this study of recall through knowledge test is only one possible way of using one possible game in the classroom. Similar studies might be conducted using other games, or testing other cognitive functions, or other uses of Minecraft Education could be explored. Some ideas for such uses are cooperative problem solving, re-enacting historical events, performing the note-taking technique in a shared world, allowing students to interact with and discuss each other's notes and so on. As the game is based around exploration and creative expression, the possibilities seem to be limitless. And as video game technologies like VR keep advancing, with increasingly immersive and comprehensive experiences becoming available, there has never been a better time to explore these possibilities.

5. Conclusion

The field of digital game-based learning is a strange one. Despite efforts to understand this subject being made already 30+ years ago, a clear narrative is yet to emerge. Looking at historical trends, a major cause for this appears to be a lack of coordination regarding research of this kind. A general lack of research also keeps getting reported. Most of the existing literature is centered around serious games, with entertainment video games receiving comparatively little attention. There could be serious untapped potential within this field.

Research on serious games reports entertainment as a big contributing factor for player immersion, and their likelihood of achieving a flow state. Research on multisensory and emotional learning provide further insight into how video games might facilitate learning.

This study explored one possible way Minecraft Education could achieve this, by having participants, Norwegian high school students, take notes within the game followed by a knowledge test to measure their recall. Due to attrition, scores on this test, along with self-

reported attitudes towards the experience were compared with only one of the two control conditions, the reading group.

No significant effect was found between study condition and recall. However, significant effects were found between enjoying the assigned task and playing the game, alongside other significant interactions. Also, worth noting is the effect between motivation to perform well and being excited about the task, which both predicted test scores. Interestingly, results suggest task enjoyment as a possible moderator of these variables. This implies the existence of an indirect effect of the study condition on recall, setting a precedent for future research.

Looking beyond the research question, the direct effects found in this study paint a picture where motivation is a strong predictor of test scores and task attitudes. They also show that students prefer reading to Minecraft-notetaking as a study technique.

Besides exploring notetaking in Minecraft, the aim of this paper was to examine where digital game-based learning is today and provide direction for future research. The findings suggest that above all else an organized effort and framework is required to properly research this topic. With that provided, there are so many potential applications of video games in educational settings. Even within Minecraft Education they appear limitless! As technological advance is at an all-time high, this should rather happen sooner than later, so the classroom can advance alongside the rest of the world.

6. References

- Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *American Psychologist, 69*(1), 66–78. https://doi.org/10.1037/a0034857
- Powers, K. L., Brooks, P. J., Aldrich, N. J., Palladino, M. A. & Alfieri, L. (2013). Effects of video-game play on information processing: A meta-analytic investigation. *Psychonomic Bulletin Review,* 20, 1055–1079). https://doi.org/10.3758/s13423-013-0418-z
- Bediou, B., Adams, D. M., Mayer, R. E., Tipton, E., Green, C. S., & Bavelier, D. (2018). Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychological Bulletin, 144*(1), 77–110. https://doi.org/10.1037/bul0000130
- Sala, G., Tatlidil, K. S. & Gobet F. (2018) Video game training does not enhance cognitive ability: A comprehensive meta-analytic investigation. *Psychological Bulletin, 144*(2), 111-139. doi: 10.1037/bul0000139
- Martinez, L., Gimenes, M. & Lambert, E. (2022). Entertainment Video Games for Academic Learning: A Systematic Review. *Journal for Educational Computing Research, 60(5)*, 1083-1109. DOI: 10.1177/07356331211053848
- Gao, F., Li, L. & Sun, Y. (2020). A systematic review of mobile game-based learning in STEM education. *Education Tech Research Dev* 68, 1791–1827 https://doi.org/10.1007/s11423-020-09787-0

- Ravyse, W.S., Seugnet Blignaut, A., Leendertz, V. & Woolner, A. (2017). Success factors for serious games to enhance learning: a systematic review. *Virtual Reality, 21*, 31–58. https://doi.org/10.1007/s10055-016-0298-4
- Hussein, M. H., Ow, S. H., Cheong, L. S., Thong, M. K. & Ebrahim, N. A. (2019). Effects of Digital Game-Based Learning on Elementary Science Learning: A Systematic Review. *IEEE Access*, 7, 62465-62478. doi: 10.1109/ACCESS.2019.2916324.
- Chuang, T. & Chen, W. (2007). Effect of Computer-Based Video Games on Children: An Experimental Study. *Educational Technology & Society, 12(2),* 1-10. https://www.jstor.org/stable/jeductechsoci.12.2.1
- Ricci, E. K., Salas, E. & Cannon-Bowers, J. A. (1996). Do Computer-Based Games Facilitate Knowledge Acquisition and Retention? *Military Psychology, 8(4),* 295-307. DOI: 10.1207/s15327876mp0804_3
- Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in cognitive sciences*, *12*(11), 411–417. https://doi.org/10.1016/j.tics.2008.07.006
- Seitz, A. R., Kim, R., & Shams, L. (2006). Sound facilitates visual learning. *Current biology*, *16*(14), 1422–1427. https://doi.org/10.1016/j.cub.2006.05.048
- The Newton Project. (2023). *About.* Sourced 10.04.2023 from https://www.newtonproject.eu/
- Tyng, C. M., Amin, H. U., Saad, M. N. M. & Malik, A. S. (2017) The Influences of Emotion on Learning and Memory. *Frontiers in Psychology, 8*, 1454-. doi: 10.3389/fpsyg.2017.01454
- Russel, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology, 39(6),* 1161-1178. https://doi.org/10.1037/h0077714

- Madan, C. R., Scott, S. M. E., & Kensinger, E. A. (2019). Positive emotion enhances association-memory. *Emotion, 19*(4), 733–740. https://doi.org/10.1037/emo0000465
- Holt, N. J., Bremner, A., Sutherland. E., Vliek, M., Passer, M., & Smith, R. (2019). Psychology: The Science of Mind and Behaviour (Fourth Edition). London: McGraw-Hill Education
- Egefeldt-Nielsen, S. (2006). Overview of research on the educational use of video games. *Nordic Journal of Digital Literacy, 1(3),* 184-214. https://doi.org/10.18261/ISSN1891-943X-2006-03-0
- Squire, K. (2003). Video games in education. *Int. J. Intell. Games Simul., 2,* 49-62. http://www.savie.ca/SAGE/Articles/1064_000_SQUIRE_2003.pdf
- Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan, 87*(2), 105-111. https://doi.org/10.1177/003172170508700205
- Nebel, S., Schneider, S. & Rey, G. D. (2015). Mining Learning and Crafting Scientific Experiments: A Literature Review on the Use of Minecraft in Education and Research. *Journal of Educational Technology & Society, 19(2),* 355-366. http://www.jstor.org/stable/jeductechsoci.19.2.355
- Mavoa, J., Carter, M. & Gibbs, M. (2017). Children and Minecraft: A survey of children's digital play. *New Media & Society, 20(9),* 3283-3303. https://doi.org/10.1177/1461444817745320
- Mumbojumbo. (2019, 24. April). *REMOVING THE WORLD: SciCraft Survival Tour FINALE!* [Video]. Yotube. https://www.youtube.com/watch?v=O12_KwtI4W4&list=PLNYE9JRimfXsLXQzhOp6_5aJP40D-Ue1i&index=5&ab_channel=MumboJumbo

- Sammyuri. (2022, 6. September). *I made Minecraft in Minecraft with redstone!* [Video]. Youtube. https://www.youtube.com/watch?v=-BP7DhHTU-I&ab_channel=sammyuri
- Ronconi, A., Veronesi, V., Mason, L., Manzione, L., Florit, E., Anmarkrud, Ø. & Bråten, I. (2022). Effects of reading medium on the processing, comprehension and calibration of adolescent readers. *Computer & Education, 185, 1-14.* https://doi.org/10.1016/j.compedu.2022.104520
- Bui, D. C., Myerson, J., & Hale, S. (2013). Note-taking with computers: Exploring alternative strategies for improved recall. *Journal of Educational Psychology, 105*(2), 299–309. https://doi.org/10.1037/a0030367
- Lesbarhetsindeks. (2020, 12. November). I *Store Norske Leksikon.* https://snl.no/lesbarhetsindeks
- Nordtømme, C. K. (2023). *Liks.* Skriftlig. https://skriftlig.no/liks/
- Csikszentmihalyi, M. (2000). *Beyond boredom and anxiety.* Jossey-Bass.
- Hemenover, S. H. & Bowman, N. D. (2018). Video games, emotion and emotion regulation: expanding the scope. *Annals of the International Communication Association, 42(2),* 125-143. https://doi.org/10.1080/23808985.2018.1442239
- Nasjonal Digital Læringsarena. (2023). *NDLA*. https://ndla.no/
- Google. (2023). Google Scholar. https://scholar.google.com/schhp?hl=en&as_sdt=0,5
- Paper Digest. (2023). https://www.paperdigest.org/aboutus/
- Elicit. (2023). https://elicit.org/
- Litmaps. (2023). https://www.litmaps.com/
- Kahoot. (2023). https://kahoot.com/
- Mojang. (2023). *Minecraft Education.* https://education.minecraft.net/en-us

7. Appendix

- **Materials & SPSS files:**

  https://drive.google.com/drive/folders/1IETp9i2fNyQVTkCm4iSO7YeMFa7TvNV0?usp=sharing

- **Distraction task:** https://create.kahoot.it/share/alvorlig-forskningskahoot/cdbd8791-2b91-4843-8a0e-01e4fcc4ebb2