博 士 論 文

Utilization of Virtual Reality in eLearning and its User Experience Evaluation
e ラーニングとユーザー体験評価におけるバーチャルリアリティの活用

Supervisor: Professor Hisashi Sato
指導教員： 佐藤尚　教授

（ 1795001 ）

Pin Hsieh/Rex Hsieh

謝ピン/謝レックス

神奈川工科大学
情報工学専攻

提出日：2023 年 1 月 11 日

Abstract

Advancement in technology has opened new ways of passing down information and the maturity of virtual reality is no different as various industry sectors embraced this technology. In recent years, the education industry training sectors especially has come to utilize Virtual Reality and related technologies, resulting in the improvement of eLearning and immersive training systems (ITS). This dissertation seeks to answer the question of whether a form of VR technology, VTubers (Virtual YouTubers) can be effectively used in the eLearning environment. VTubers are essentially broadcasters who digitally transformed themselves into anime styled avatars while sometimes using voice transform devices to mask their original audio, therefore hiding their original self from target audiences completely. The research team has also came up with a variety of evaluation methods to better measure the effectiveness of virtual reality contents.

論文要旨

仮想現実技術の進歩により、さまざまな問題に取り組むさまざまな VR アプリケーションが増えています。研究者が仮想現実を利用しようとしたこれらの問題の中で、教育、より具体的には e ラーニングの領域があります。この論文は、VR テクノロジーの一種である VTubers(バーチャル YouTuber) が eLearning 環境で効果的に使用できるかどうかという問題に答えようとしています。VTuber は本質的に、デジタルで自分自身をアニメスタイルのアバターに変換し、音声変換デバイスを使用して元の音声をマスクし、対象の視聴者から元の自分を完全に隠した放送局です。研究チームは、バーチャル リアリティ コンテンツの効果をより適切に測定するためのさまざまな評価方法も考案しました。

目次

第 1 章

Introduction

1.1 ITS, eLearning, and VTuber

The eLearning industry is a rapidly developing industry and is expected to continue growing due to the increasing focus on remote education. According to Docebo's report on eLearning trends published in 2016 [41], the eLearning market in 2015 is worth 165 billion USD and it is expected to grow to over 240 billion USD by 2023 partially due to the eLearning users developing an appreciation for animated and engaging contents as well as the proliferated usage of online and mobile devices. In short, by integrating creative contents into the education system, students get to absorb new information through fun and meaningful ways which not only made learning easier but more efficient.

One of the newer inventions as a result of virtual reality gaining popularity is VTuber also known as Virtual YouTuber. In 2018 it was estimated the number of VTubers is in the 4,000s, which is a significant increase from the 2,000s in 2017. Some of the popular virtual Youtuber applications currently in use include *upd8* and *Reality*, each enjoying over a million users. Currently the most popular VTuber star: Kizuna AI enjoyed over close to 3 million followers and has generated a variety of contents ranging from tutorials to talk shows involving with other VTubers and even real people.

1.2 Motivation and Background

The very first section of this dissertation introduces a project name Real Baby - Real Family, which serves as the inspiration and motivation for the second research: Mas-

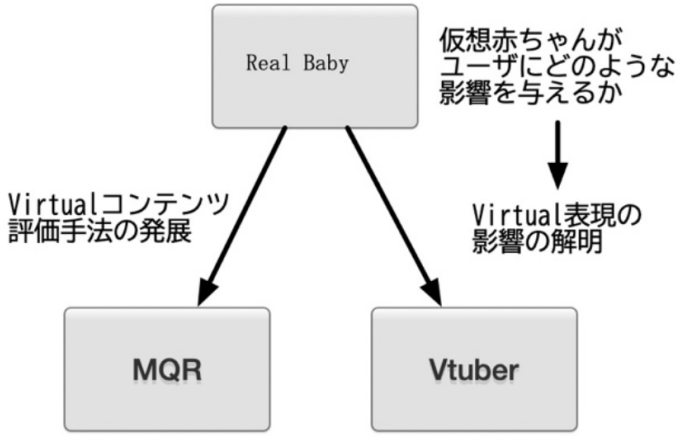

図 1.1 Research 1: Real Baby's virtual reality nature and how it impacted exhibition attendees served as the inspiration for MasQueRade (MQR) and VTuber styled eLearning Videos.

QueRade, a VR experience evaluation system utilizing QR code and third research: VTuber styled eLearning Videos. Real Baby - Real Family is a Virtual Reality based baby nursery simulator and an immersive training system that allows users to experience what it's like to raise their own children.

The next project to be introduced is a VR experience evaluation system utilizing QR code that was developed drawing lessons from Real Baby - Real Family's exhibition procedure. This concept was inspired by the long waiting lines, also known as queue, of attendees standing by exhibition booths waiting to experience a single experience and the inconvenience this caused to exhibitors, attendees, and people trying to walk around the queue. Although queue managing applications [43] have been created to better aid exhibitors and organizers managing attendees, the research team believed it is more efficient to decrease queue size and free up attendees to participate in other activities. MQR was developed so it could store attendee's information and their waiting number so as to allow them to engage in other activities.

Furthermore, the experience can be made more complicated if the exhibitors decided to survey opinions from attendees, which can can easily lead to severe delay in everyone's schedule and requiring additional manpower to hand out and gather ques-

tionnaires. Other issues that can arise from conducting survey during events to mention requiring additional manpower to conduct the survey. One other problem traditional paper-based questionnaire has is the amount of waste produced afterwards. According to Japan Environmental Sanitation Center , the annual waste management expense of mixed paper is 681,718 yen [42]. The United States Environmental Protection Agency [44] pointed out in 2018, 67,390 thousand tons of papers were produced and amongst them, 45,970 thousand tons were recycled. While the MQR was not developed with the intention of saving on paper cycling expenses, the research team still hoped the usage of QR code based evaluation system can help decrease paper waste.

The next part of this paper described the research team's attempt at integrating VTuber styled videos into an computer science university curriculum. Due to the popularity of VTuber and wide spread cultural acceptance of user generated anime-styled contents in Japan, the research team believe the supplementation of VTuber-like videos in traditional face to face university lectures will help boost interest and grade performance of students by providing additional way for students to absorb information. Furthermore, due to the many different types of VTuber anime avatars, the most common of which being female-styled avatars, this research believed it is essential to test different forms of VTuber avatars and record students' video watch duration, academic performance, and view frequency to analyze strength and weakness of each avatar.

1.3 Dissertation Structure

This dissertation is separated into the 2 major research projects: *MasQueRade* (MQR) and VTuber eLearning program along with related researches. While each containing their unique attributes, are all tied together by their user engagement focused user experience approach. Following these 2 main researches are the minor researches the research team has done followed by the conclusion.

Chapter 2 introduces Real Baby – Real Family, a project started by a team of senior students which won the Virtual Reality International Contest in 2016 and was exhibited in Laval Virtual 2017 and SIGGRAPH 2017. The author's contribution to this project

and results of international exhibitions as well as its influence on later project are described in detail.

Chapter 3 is dedicated to MQR which aims at improving the ways virtual reality exhibitors collect data by assigning QR codes to attendees and letting them answer survey questions on their free time. The system design, data gathered, and user feedback from several exhibitions and conferences, including the Anime Expo 2017 and SIGGRAPH 2017 are covered.

Chapter 4 describes the VTuber eLearning video system. This research focuses on improving student participation and academic grade performance in university programming course by introducing eLearning videos featuring anime styled avatars of different genders and audios. Student impression score and view count of each video as well as academic performance are covered.

Chapter 5 serves as the conclusion of this paper by summarizing all of the researches and findings. This chapter concludes by talking about the future for these projects.

第 2 章

Real Baby – Real Family

The *Real Baby - Real Family Age Controllable VR Avatar from 2D Facial Images* is a project aimed at creating digital facial features for different avatars from any 2D facial images. By pulling specific facial features from photographs, analyzing them, and then merging the obtained data together, our team has successfully recreated digital baby avatars whose faces closely resembled that of subjects'. Furthermore, we can generate digital avatars not just from real-life photos, but also from illustrations and even a mixture of photos and drawings.

2.1 Related Works

Neuro-Baby [14], *Babybot* [15], *Infanoid* [16], *Kismet* [17] [17], *Cog* [18], *YOTARO* [19] are some of the baby nursery simulators conducted by other researchers. Amongst these researches, *YOTARO* can be considered as a work that emulates a realistic baby. *YOTARO* is a baby robot that can simulate a runny nose and many different emotions. However, because *YOTARO*'s face is that of an illustration and not a real-life portrait, it looks as if the baby is not related to the players. Our project generates a baby from portraits of heterosexual or homosexual couples playing it thus players get a baby looking just like the players. While *YOTARO* can be considered a predecessor that represents a realistic VR baby robot, its face which is composed of an illustration, makes it look like a fictional character or a baby belonging to another person.

Face2Face [20] is another project we look into. It uses monocular facial reenactment in real-time to manipulate the target video into acting out the facial motions of the source. This project deals with real time video manipulation and not still image ma-

図 2.1 From left to right: (1)68 corresponding points, (2)Screen-capture of actual gameplay of player holding onto VR baby, (3)User wearing HMD in real-life feeding virtual baby milk bottle using HTC Vive controller

nipulation; however, and it can only manipulate the facial motion and not the age progression of the portraits.

2.2 Design

This VR projected realized a virtual family enabling the players to communicate with a baby through haptic, visual, and audio feedback while wearing a head-mounted display (HMD). Figure 2.1 shows (1)the 68 corresponding points of virtual baby's face used to integrate user's image into virtual baby face, (2)VR experience in-game capture, and (3)what users looked like during experience.

According to the proposed hypothesis, creating a believable baby resembling the players will make people more emotionally invested with the experience. Following this theory, we created a virtual baby generated from players' photos. The most important part of our project; however, is that we constructed a physical baby that can interact with the audience. Real Baby - Real Family is unique for allowing players to (1)hug and interact with the baby physically, (2)creating a baby looking just like the players, (3)contains audio components.

2.2.1 Holdable robot device

In order to hug a baby doll without experiencing discomfort while wearing the HMD, there is a need for high precision position tracking of the doll. However, this is hindered by image recognition ability of an one eyed camera which slows down the position tracking speed. Our proposed method uses multi-point image sensor implemented in the controller of HTC Vive to enable high precision position tracking. This synchronizes

the physical and virtual movement of the baby and allows players to hug the baby without getting visually disoriented.

2.2.2 Visual Face Generator

Below are three characteristics of our baby face generator: (1) Generating a face fitting the players' skin tone, (2) creating the baby's face from multiple photos, (3) inverse the age of the players to fit that of the baby.

2.2.3 Average Baby Face Generation

It determines the skin tone of the baby by averaging the skin tone of the two photographs. Our system also morphs the colors and shapes of the faces using *OpenCV*. Lastly it collects 16 baby face images using Japanese Google Image Search.

2.2.4 Get Face Landmark Index

In order to morph the many photographs, baby images, corresponding points in the feature points, it is necessary to obtain the index. In the proposed method, we use the *Dlib* of open source library, from the results that have been learned by the data set iBUG-300-W, to get the points of each part of the face. The eyes, noses, mouths, eyebrows, feature points, and the 68 points with index composed of contour, are all morphed automatically. Here it can be seen that morphing is concentrated in the mouth and even parts of the face.

2.2.5 Inverse Age Progression

Inverse age progression is a process of generating an average from the feature obtained by the face image and (2), without collapse baby portraits obtained in (1). The implementation uses a *Dlib* and *OpenCV*.

The obtaining of feature points is performed at *Dlib*, which performs image generation by passing the coordinate data of the feature points to the *OpenCV* side. *Dlib* automatically sorts the feature points obtained, treating them as landmarks. Its index is unchanged in all of the facial image. By setting the contribution ratio with respect

to the index number, it is possible to perform processing for each selective parts. The color components are produced by (1), the present process is a concept that only features the specified part that were inherited. If carried out well as a deformation of the 3D modeling, it should also be considered, such as contour around the chin. The generated face image is used by Live.

Inverse age progression is the process of generating a final image from the averaged age of portraits at Get Face Landmark Index. This is all done without collapsing baby portraits obtained in Average Baby Face Generation. The implementation uses a *Dlib* and *OpenCV*. It obtains feature points performed at *Dlib*, performs image generation by passing the coordinate data of the feature points *OpenCV* side. *Dlib* automatically sorts the feature points obtained and treated them as a landmark. Its index is unchanged in all of the facial images. By setting the contribution ratio with respect to the index number, it is possible to process each selective parts. The color components are produced by Average Baby Face Generation, the present process is a concept that only features the specified parts inherited. If carried out well as deformation of 3D modeling, contour around the chin and other aspects should also be considered. The generated facial image is used by Live.

2.2.6 Dynamic texture generation

In generated images only pasted to the 3D model has a large discomfort, it is necessary to generate a facial expression that varies vividly dynamic. Therefore, we use the "image deformation software" "Live" for character animation, to prepare the animation set in advance, and real-time generated in the VR video space in Unity. In creating the model data of Live, the image cut out the part of the eyes and the mouth of the "average face of the baby face image" as a template.

2.2.7 Display of Haptics Synchronized with Voice

When holding the baby robot doll, the baby robot is strengthening the interaction by vibro of Vibro transducer Vp2 (Vp210) which was mounted in the baby doll. This vibro is controlled by voice. Only by passing the band-pass filter using a Fourier transform

effective frequency band (5Hz 200Hz) and extracted as the vibration data onto an output audio can it perform real-time tactile presentation through HMD attendant controller (Vive Controller). However, since it's the only vibrator that is built in the standard Vive controller, we felt its electric current output which vibrates the baby doll was insufficient after many exhibitions. To fix this problem we mounted the Vp210 onto Vive controller to improve the oscillating capability.

2.3 Conclusion

By synchronizing the audio, visual, and haptic feedback of the physical and virtual baby, we are able to realize a believable baby nursing experience. Furthermore, we succeeded in creating a baby portrait that borrows features from the players' photos and created a virtual baby capable of fostering deeper connections with the players. A nursing simulator child care technique was also tested in this project and the result showed this project has the potential to become a VR nursing simulator. In the future we are hoping to create more VR entertainment systems featuring a complete VR world and real life physical objects such as 2D facial pictures and baby dolls.

第 3 章

MasQueRade

3.1 Introduction

This section describes the development and function of *MasQueRade*, an on-site VR application evaluation system that stores surveys inside QR codes and can be attached to a variety of devices: VR hygiene mask, name card, and cosplay devices. By attaching QR codes to devices commonly wore or carried by attendees, the research team realized a portable survey that allows users to evaluate the VR experience at their own leisure. Figure 3.1 showed the SIGGRAPH 2017 version of MQR with the QR codes attached to hygiene masks.

This research is motivated by the increase in number of VR applications and how almost every single digital entertainment company is investing heavily in VR systems. This increase in VR products demands improvement of evaluation methods that can speed up and free up attendees from the tedious task of waiting in line. Furthermore, some VR experiences do not look especially impressive from an onlooker's perspective but managed to impress players during experience. In cases like these, a before and after questionnaire is needed to measure the user's initial impression and actual VR walk-through impression. *MasQueRade* was developed to lessen the work required for

図 3.1　MQR SIGGRAPH 2017 version

conducting extensive survey in a time efficient way.

3.2 Past Researches

Aoto and Ohkura aim at objectively realizing an evaluating system by quantifying the feeling of excitement that is triggered when one feels something by making use of biological signals. In their study, an evaluation method for detecting comfort and discomfort via questionnaire and determining the degree of discomfort through biological signals by simultaneously conducting a questionnaire was proposed [1] [2]. While Bio-metrical methods have a chance to build objective evaluation, it is complicated and uncomfortable and should be avoided in an exhibition where attendees' goal is to enjoy themselves.

A number of human based evaluation methods have also been proposed which utilizes groups of evaluators to give lengthy play-tests before giving the virtual reality system scores. Heuristic method [34] [40] is one such method that have the evaluators familiarize themselves with the VE/VR system, carry out a set of tasks, list problems encountered, before classifying the problem. The issue with this approach being it uses experienced users and not everyday folks. MQR aimed at allowing even average users to give inputs regarding the VR system.

Other VE evaluation methods include the Cognitive Walkthrough [36] [37] which evaluates user interface by evaluating the effectiveness of interface at helping users perform tasks, Formative Evaluation [38] which places users in task based scenarios to identify usability problems, Post-Hoc Questionnaire [39], Interview, and Summative or Comparative Evaluation which uses statistical approach to compare and contrast two or more UI designs [35]. MQR is a system that evaluates the user's initial impression of the VE/VR and then conducts the same questionnaire once more after the experience has been concluded, thereby measuring the before and after impression of a system and compare user impression against user evaluation score.

3.3 Design Philosophy

MQR is mainly mounted onto a sanitary mask made of non-woven fabric and used in conjunction with Head Mounted Display. It is composed of (1) Player ID, (2) QR Code, and (3) NFC tag.

Players and or operators can access the survey site by coded URL on MQR. Their answers are sent to a server which is then saved to a database with Application Programming Interface (API). Since operators and VR developers can access this data containing properties and feedback of players, it is easier for them to improve their entertainment system. Players on the other hand enjoy a smoother survey method.

3.3.1 Before Playing

In VR events, most entertainment systems can be finished in less than 5 minutes. This being said, the amount of time spent putting on gears, listening to explanations, and finally answering questions make switching between players difficult and time consuming. With MQR, users can scan the QR code on MQR, answer the questionnaires, and input any information required to experience the VR application beforehand therefore saving time. In addition, once players have answered all the surveys, their MQR will automatically mark down their time stamp meaning the users can go to experience other activities and return at a later time, therefore decrease unnecessary time spent standing around waiting for their turn.

3.3.2 Experience Adjustment based on Feedback

During the gameplay, the proposed method also contributes to the dynamic narrative design. By recording feedback generated by past players, this system can help developers change the difficulty to appropriate levels or, should the same player choose to experience the same activity once more, adjust the difficulty based on their previous feedback.

3.3.3 After Playing

Conventional surveys record players' opinions after each experience and requires space and manpower, both of which lacking in popular events. Furthermore, most Virtual Reality systems use infrared to sense the posture of the users. By conducting interview at the booth, it is possible to accidentally block the infrared and causing systems to malfunction. By using this system, players can answer questions without space and time constraints. For example, they can answer the survey while on their way back home. Players can also use the extra time to meet new people and enjoy more exhibitions. During Anime Expo 2017, employees for an application named Manga Generator was able to use MQR to obtain vital information regarding which manga stripes were the most popular, therefore allowing team members to improve upon the system after the exhibition.

3.4 Implementation

MQR was designed using MySQL and PHP. Users can gain access to the website by either using their smartphone or tablet to scan the QR code, which consists of an url followed by unique ID named the MQID assigned to individual QR codes. Upon scanning the code, users will be taken to the Landing Page which consists of links to the survey pages for different applications as well as the before and after questionnaire for the different applications.

Before the user can proceed onto the questionnaires; however, they must first agree to the user agreement form which, depends on the exhibition, may differ in content. For instance, since the European Union (EU) requires all websites that use cookies to seek user agreement, the MQR system in France specifically asks users to allow for cookie. Clicking on the "agree" button will allow users to proceed to the questionnaire while clicking cancel agreement form will log them out of the system.

Upon clicking on the survey button, the system automatically records the exact time the survey started. Likewise, once the users have finished the survey, the system will also record the exact time survey finished. This is done in an effort to record how

certain users are sure of their answers which can be used to judge the definitive value of their answers.

3.5 Exhibition Results

MasQueRade was tested in several international exhibitions, and utilized on numerous VR systems, gathering data regarding user impression. This section documents data analysis and functionality of MQR in large public events with hundreds of visitors.

3.5.1 Anime Expo Japan VR Festival 2017 Results

MQR was utilized in the 2017 Los Angeles Anime Expo which took place from June 28th, 2017 to July 4th, 2017. This event was the largest anime exposition in North America and attracted 115,726 non-repeating visitors. Within the exhibition was an event called the Japan Character VR Festival hosted by Shueisha featuring 4 VR applications created by Shueisha and 7 VR applications created by outside companies. This was when MQR was utilized to measure which applications have been played and other general information such as which VR equipment each users own.

Here is a list of the questions attendees of Japan Character VR Festival answered before they experienced the festival.

1. Please tell us your gender.
2. Please tell us your birth year.
3. Did you have any prior experience playing with VR system? Please list all VR systems that you have experienced so far.
4. Please select all VR systems you own.
5. (Your impression before playing) When you entered the booth, did any of the VR systems attracted your attention? Please rate how interesting the systems looked from 1 to 5 with 1 being least interesting and 4 being most interesting.
6. Where have you experienced VR systems before? Please select all that apply.
7. What is your purpose for visiting Anime Expo?
8. What is your reason for visiting the "Japan Character VR"?
9. Please tell us the VR title you want to play the most.

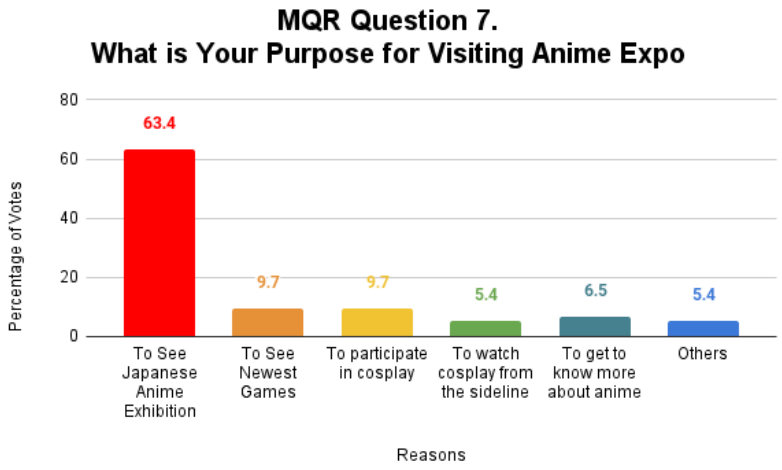

図 3.2　Poll result of Question 7: “What is Your Purpose for Visiting Anime Expo”

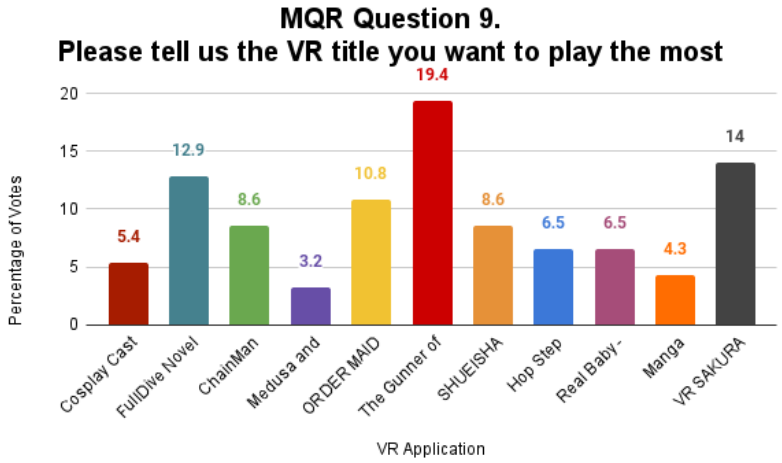

図 3.3　Poll result of Question 9: “Please tell us the VR title you want to play the most”

The research team obtained a total of 93 replies from attendees showing their impression of the exhibition. Figure 3.2 shows their answers to question 7: “What is your purpose for visiting Anime Expo?”

A majority of attendees went in order “To See Japanese Anime Exhibition”. With “To Participate in cosplay” and “to See Newest Games” the next 2 most voted answers. Figure 3.3 shows the answer to question 9: “Please tell us the VR title you want to play the most”.

Figure 3.4 shows the answer to Question 5, the initial impression to the 11 VR titles

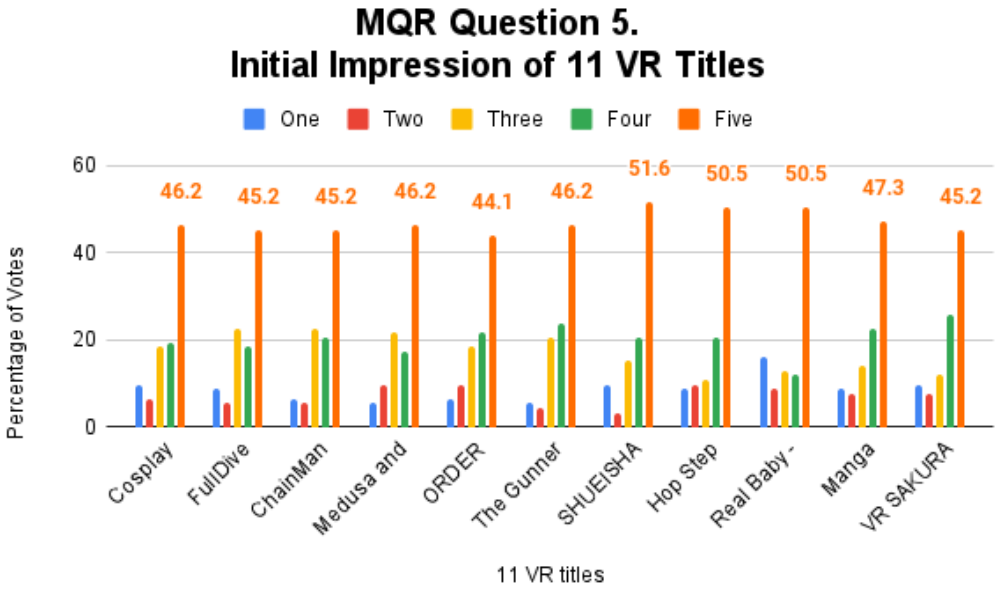

図 3.4　Poll result of Question 5: "Initial Impression of 11 VR Applications". Score goes from 1 - 5 with 1 being the lowest and 5 being the highest. Score displayed in percentage.

exhibited at the Anime Expo 2017. It can be clearly seem from the graph that the majority of the attendees viewed most applications favorably and rated them as 5 with less than 10% of attendees rating their impression of any VR application a 1.

After the attendees have experienced the VR applications, they were once again asked to take a survey regarding their impressions for the event. For this survey there are 13 questions as well as 2 general information questions inquiring about their gender and age. The 13 questions are:

1. Did you have any prior experience playing with VR system? Please list all VR systems that you have experienced so far.
2. Where have you experienced VR systems before? Please select all that apply.
3. What is your purpose for visiting Anime Expo?
4. What is your reason for visiting the "Japan Character VR"?
5. (Your impression before playing) When you entered the booth, did any of the VR systems attracted your attention? Please rate how interesting the systems looked from 1 to 5 with 1 being the least interesting and 5 being most interesting.

6. (Your impression after playing) Since you are only allowed to play six systems, please rate the VR experiences you have played from 1 to 5 with 1 being not fun at all and 5 being very fun.
7. Please tell us which of the systems you didn't get to try but would like to try in the future.
8. Please choose all VR systems you would be willing to buy if published.
9. Please select all VR systems you own.
10. Please tell us the VR title you want to play the most.
11. While playing the VR system, did you experience any discomforts? Please select all that applies.
12. What is your reason for not using the Head Mounted Display (HMD)? Please select all that apply.
13. Please tell us any other systems you would like to try.

These surveys are answered by 10 attendees with 9 of them being male and 1 being female. Figure 3.5 is the graph for Q1., showcasing that Oculus Rift is the VR platform most people have experienced followed by Gear VR, HTV Vive, PlayStation VR, and other unspecified VR. No attendee has indicated that they have experience with Mobile VR and Google DayDream View. Furthermore, 4 of the attendees indicated they have no experience with any VR platform.

Figure 3.6 displays the results for Q2., with most attendees indicated that they have experienced VR at the game center followed by home and friend's home, and lastly school and VR center. 4 of the attendees once again indicated they have not experienced VR system anywhere.

Question 5 asks for the attendees to rate the initial impression for the 11 VR contents. Of these VR applications, *Medusa* is the most highly rated application with a mean score of 4.3 followed by *Full Dive Novel* and *VR Sakura* both at 4.1, *Chain Man* and *Order Maid VR* and *The Gunner of Dragon* at 4, *Cosplay Cast* and *Real Baby – Real Family* at 3.8, *Hop Step Sing* and *Manga Generator JUMP VR* at 3.7, and finally *Shueisha Jump VR* at 3.6. Figure 3.7 showcases the results for Q5.

Question 6 asks for the attendees to rate the VR content they have experienced

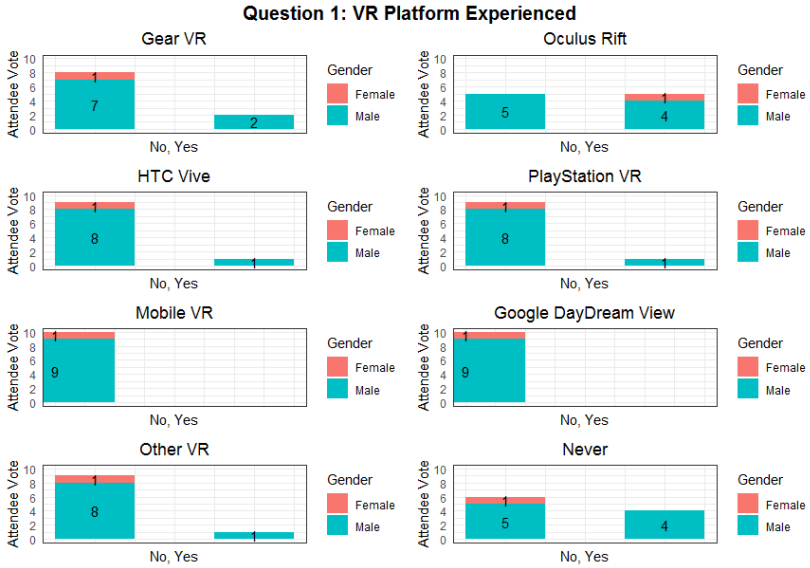

図 3.5　Anime Expo Japan VR Festival After Played Q1 Response

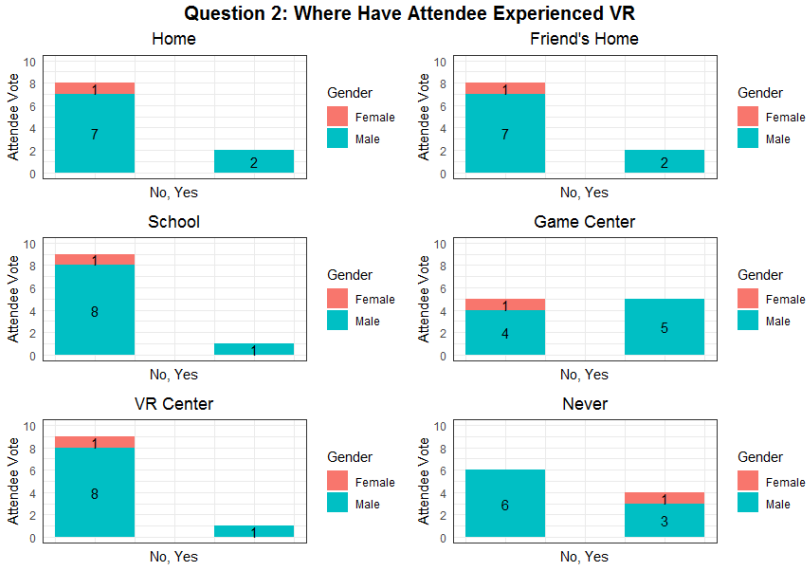

図 3.6　Anime Expo Japan VR Festival After Played Q2 Response

from a scale of 1 to 5 in an attempt to compare the attendees' initial impression with after play score. Of these VR applications, *Chain Man* is now the most highly rated application at 4.2 followed by *The Gunner of Dragon* at 4.1, *Cosplay Cast* and *Medusa* at 4, *Full Dive Novel* and *Manga Generator* at 3.9, *Shueisha Jump VR* and *Hop Step Sing* and *Real Baby – Real Family* at 3.8, *VR Sakura* at 3.7, and finally *Order Maid VR*

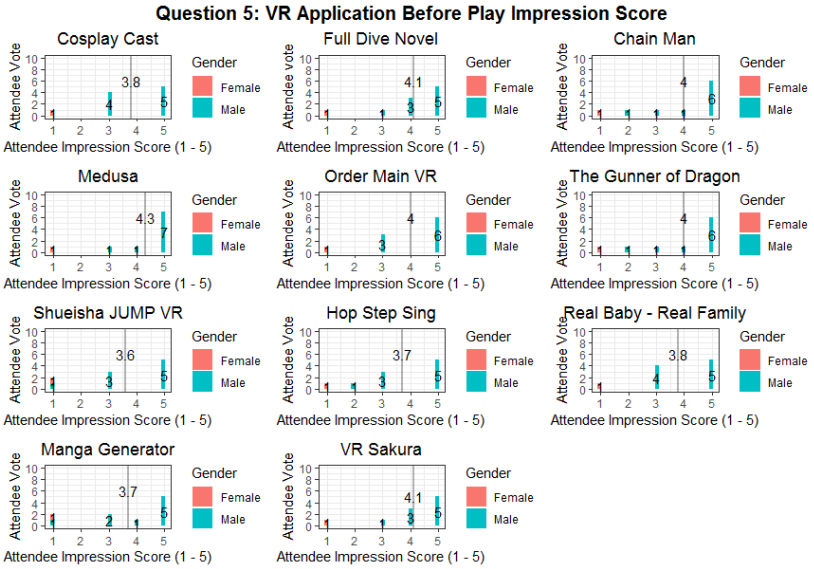

図 3.7　Anime Expo Japan VR Festival After Played Q5 Response

at 3.6. In ranked order from most favorable change to least favorable change between initial impression and after play score is as follows: *Cosplay Cast*: 0.2, *Chain Man*: 0.2, *Shueisha Jump VR*: 0.2, *Manga Generator JUMP VR*: 0.2, *Hop Step Sing*: 0.1, *Real Baby – Real Family*: 0, *The Gunner of Dragon*: -0.1, *Full Dive Novel*: -0.2, *Medusa*: -0.3, *Order Main VR*: -0.4, *VR Sakura*: -0.4. We can see from these data that while the applications that increased in score and those decrease in score are exactly the same both at 5, the amount of increase is locked at 0.2 while the amount of decrease is double that amount at -0.4. Despite of this, it cannot be argued any application did poorly as all still scored a 3.6 out of 5 or above in both initial impression and after play score. The score for Q6. Is shown in Figure 3.8.

Question 7 asks the question which of the VR applications did the attendees not get to play but would like to try out in the future. Here in ranked order from most want to play to least is as follows: *Chain Man*: 6, *Cosplay Cast*: 5, *Full Dive Novel*: 5, *Order Maid VR*: 4, *The Gunner of Dragon*: 4, *Hop Step Sing*: 4, *Medusa*: 3, *Shueisha Jump VR*: 3, *Manga Generator JUMP VR*: 3, *VR Sakura*: 3, *Real Baby – Real Family*: 2, and finally one attendee voted he wanted to play none of the applications. The 1 female participant voted that the only VR application she didn't get to play but would want to is Cosplay Cast. The visual representation for Q7 can be seem in Figure 3.9.

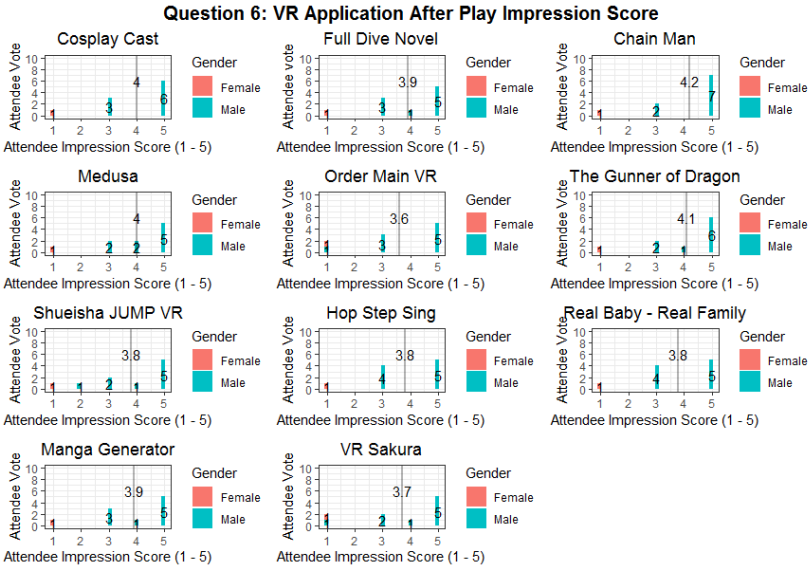

図 3.8　Anime Expo Japan VR Festival After Played Q6 Response

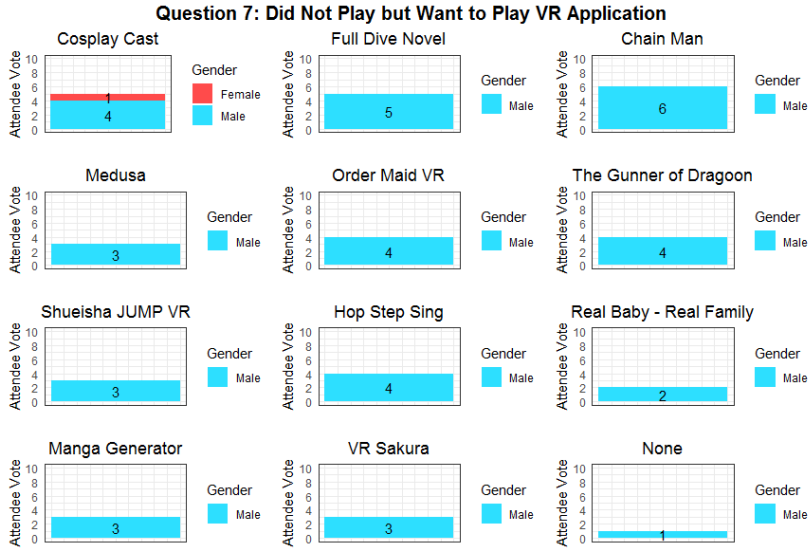

図 3.9　Anime Expo Japan VR Festival After Played Q7 Response

Question 8 asks the attendees which VR application are they willing to buy if published and according to the rank of popularity is as follows: *Cosplay Cast*: 4, *Full Dive Novel*: 4, *Chain Man*: 4, *Medusa*: 4, *Order Maid VR*: 4, *The Gunner of Dragon*: 4, *Shueisha Jump VR*: 4, *Manga Generator JUMP VR*: 4, *Hop Step Sing*: 3, *VR Sakura*: 3, *Real Baby – Real Family*: 2. Once again one voter said they would not buy any of

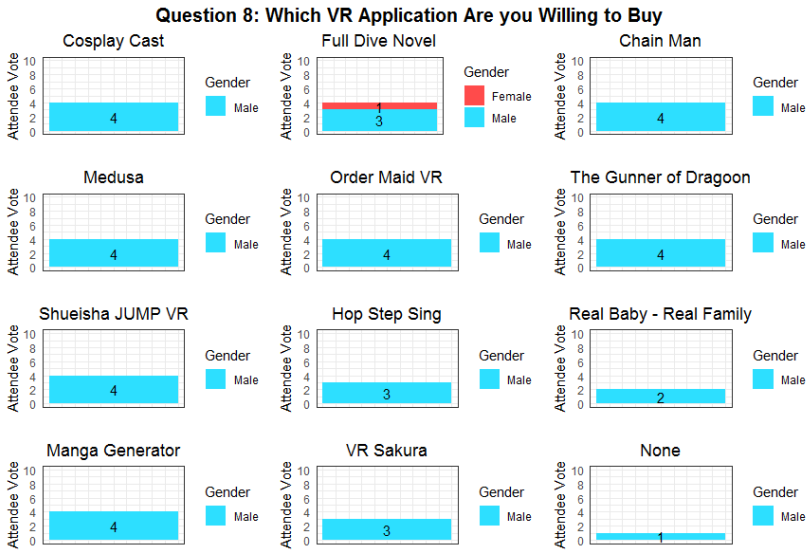

図 3.10　Anime Expo Japan VR Festival After Played Q8 Response

the applications. This shows the marketability of the different VR applications and can be used to give developers a clear picture of which applications to develop further. Figure 3.10 shows results for Q8.

Question 9 asks if the attendees owned any of the VR platforms and to some degree is a repeat of question 1. Here with the exception of HTC Vive which receives 2 votes, all other VR platforms got 0 vote. It is worth pointing out that there is contradiction between results from question 9 and question 1; however, as only 1 person indicated they have experienced HTC Vive in question 1 but in question 9, 2 attendees said they owned HTC Vive. This contradiction can only be concluded in 2 ways: 1. The female attendee clicked on the wrong button or 2. The female attendee once owned HTC Vive but never used it herself (the research team does not rule out the possibility she may have bought it for her family members or friends). Still the results showed a vast difference between number of attendees having experienced VR platform. Figure 3.11 shows the results for Q9.

Question 11 and 12 concerns about the comfort of attendees during experience with Q11 shown in Figure 3.12 inquiring if attendees have experienced any discomfort when experiencing the VR systems and 5 attendees indicated they were wearing glasses, 1 experienced strabismus (cross eye), while another attendee experienced dizzy during

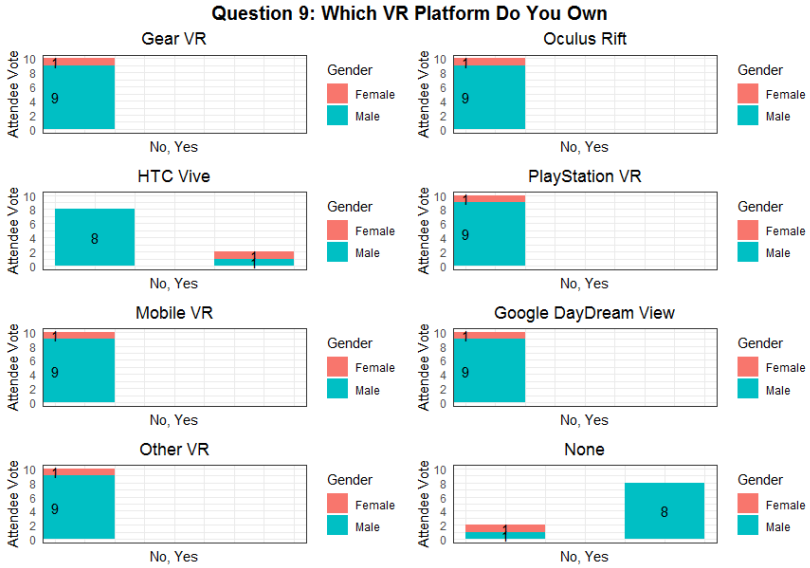

図 3.11　Anime Expo Japan VR Festival After Played Q9 Response

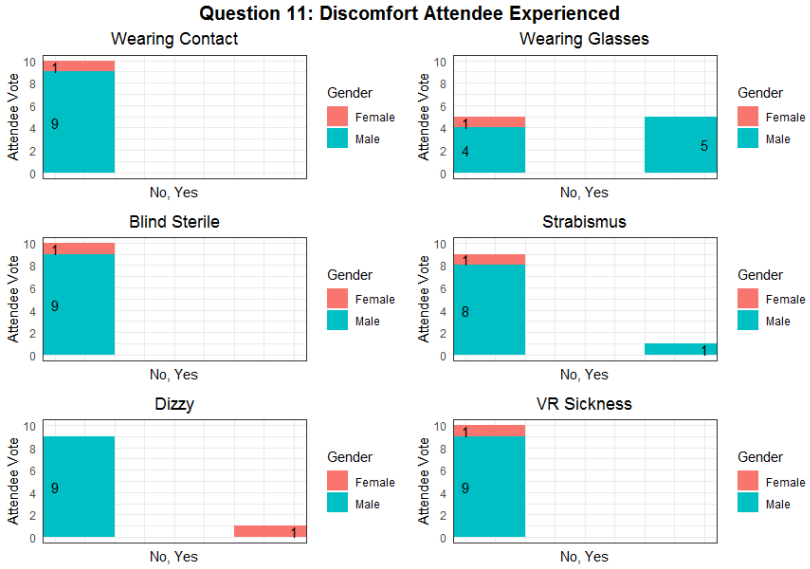

図 3.12　Anime Expo Japan VR Festival After Played Q11 Response

experience. Q12 as shown in Figure 3.13 inquired if attendees did not wear HMD, what had caused them to make the decision. Here no attendee indicated it was because of makeup, 1 female indicated it is because of her hair, 1 male indicated it is not hygienic, and 3 other attendees voted other with 2 claiming they did worn HMD and 1 saying "the person in front was using it".

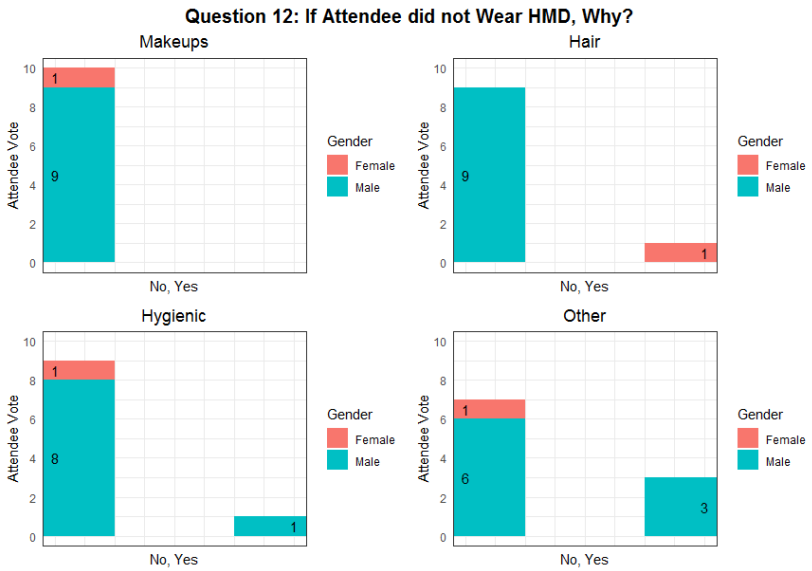

図 3.13　Anime Expo Japan VR Festival After Played Q12 Response

MasQueRade is able to evaluate not only the technical background of each attendant, but also compare and contrast their impressions for each VR application as well as marketing research such as willingness to purchase certain VR applications and their replayability.

3.5.2　Anime Expo 2017 Manga Generator Results

Aside from general attendee impressions regarding contents in an exhibition, MQR can also be used to measure specific applications and the various scenarios of such application. One of the contents exhibited at Japan VR Festival in Anime Expo 2017 named *Manga Generator JUMP VR*, for instance, was given further researches by having questionnaire specifically designed for its 4 manga scenarios.

These 4 manga scenarios, only playable at Anime Expo 2017, were all specifically designed for the event. By collaborating with Shueisha, the research team was given the permission to design 4 interactive manga scenarios utilizing 2 of Shueisha's most popular franchises at that time: *One Piece* and *Naruto*. The 4 playable interactive mangas were given the following names: *Naruto: Revolution*, *Naruto: Bring it On*, *One Piece: Mihawk*, and *One Piece: Nest of Neptunians*.

MasQueRade was implemented onto *Manga Generator JUMP VR* just like the other

VR applications. The research team was able to obtain 10 surveys from players, all of whom having filled out the important questions except, in some instances, their nationalities or additional feedback. Below is a list of survey questions:

1. Where do you live?
2. Which one of the following scenarios did you experienced?
3. How much did you enjoyed this? 1 means barely enjoyed at all and 5 means enjoyed to the maximum.
4. Do you know the specific manga scenario you have just played?
5. Did you see the original piece before you make your pose?
6. Did you strike a pose similar to the reference pose?
7. Do you like your captured poses?
8. Do you want to play this experience again?
9. How do you feel about the amount of time given to experience this system?
10. Please write down any other manga titles that you wanted to play using Manga Generator?
11. Please give us any other comments or suggestions you have!

Question 1: "Where do you live", refers to the nationality of the players. 8 survey takers answered that they are from the U.S.A. while the last 2 chose not to disclose their nationalities.

Question 2 asked the participants which of the 4 scenarios they chose. Of the 10 participants who answered this question, 4 participants answered they chose *Naruto: Revolution*, 3 picked *Naruto: Bring it On*, 3 picked *One Piece: Mihawk*, while none picked *One Piece: Nest of Neptunians*.

Question 3 allowed players to pick from a scale of 1 to 5, which represents how much they enjoyed the experience. All but 1 chose 5, which stood for they enjoyed the experience to the maximum while the remaining participant picked 4. This gave the average score of *Manga Generator JUMP VR* a 4.8. It is worth noting that the one who rated this system 4 played *One Piece: Mihawk* and later reported that he/she was unable to download Manga from the QR code which may have contributed to his

dissatisfaction.

Question 4 asked if the player is aware of the story behind the manga they have just played. Surprisingly, despite of all 4 manga stripes from world famous manga titles, 4 survey takers answered "No, I don't know.". Of these 4, 1 played *Naruto: Revolution*, 2 played *Naruto: Bring it On*, and 1 played *One Piece: Mihawk*. Considering how this is an Anime Convention and event attendees are expected to be anime hobbyist at the very least, the answers recorded were highly surprising.

Question 5 asked: Did you see the original piece before you make your pose? 8 survey takers answered "Yes" while 2 answered "No". It is worth noting the 2 players who answered "No" in question 5 also answered "No, I don't know." in Question 4. This, the research team believes, is possibly due to the poor phrasing of Question 5 because "original piece" can mean either the original manga stripe or the reference pose of the original manga characters. The original question was meant to ask if they like having a reference pose. Therefore, it is best to disregard answers from Question 5.

Question 6 asked if the players struck a pose similar to the reference pose. The response to this question is an even split with 5 answering "Yes" and 5 answering "No". The response to this answer is hardly surprising as most players chose to create their own pose or testing the limits of *Manga Generator JUMP VR's* motion capture system.

Question 7 asked if the players liked the poses they made. All players answered "Yes" for this question.

Question 8 asked: "Do you want to play this experience again?" and generated all "Yes" as answers.

Question 9 questioned the players how they felt about the amount of time given to experience this system. 6 participants answered the time was "Just Right", 2 answered it was "A bit short", 2 answered it was "Way too short" and none answered "A bit long" or "Way too long". Interestingly, the 2 players who answered "A bit short" played *One Piece: Mihawk* while 1 of the player who answered "Way too short" played One Piece: Mihawk and the other one played *Naruto: Revolution*. In summary all who played *One Piece: Mihawk* expressed dissatisfaction to some degree in regards to

experience time for *One Piece: Mihawk* while none expressed that *Naruto: Bring it On* was too short.

Question 10 asked the players what other mangas they were interested in playing. *One Punch Man* got 2 mentions while Any Shojo manga, *Teacher Onizuka*, *Haikyuu*, and *Full Metal Alchemist* got 1 mention. Aside from these recorded in the survey, one player in the exhibition asked if there are any *Fairy Tail* manga available.

Question 11 was the comment section allowing for users to input any comments they desire. All but 1 comment was positive and the negative one expressed disappoint in unable to download the manga stripes.

Judging from the data, the research team could conclude that the majority of the players were very satisfied with the experience. The lack of knowledge for the original manga story does not affect the overall fun factor of this experience, as despite half answering they do not know the scenario, all but 1 player gave a 5 for how much they enjoyed the experience and all liked the poses they made as well as expressing the desire to retry the game. The data also supports the inclusion of a Fade out mechanism for the original manga character, as half the players struck a pose not similar to the reference pose. This meant having a reference pose only serves as a guideline but does not take away the creativity of the players. Before the incorporation of fade out, some players would simply stand in front of Kinect not sure what to do and this new introduction of Fade out helps solve that problem. As for the time given to experience this system, more data need to be obtained before a conclusion can be made as all 3 players who expressed playtime is too short plays the longer manga with 3 playable panels and not the shorter 2 playable panel only *Naruto: Bring it on.*

3.5.3 SIGGRAPH 2017 MQR Results

MasQueRade was implemented during SIGGRAPH 2017 exhibition alongside with Real Baby Real Family, which will be further discussed in chapter 5 of this paper. To briefly summarize the Real Baby Real Family project (abbreviated to Real Baby) here so as to better discuss the contribution of MQR to this project, Real Baby is a Virtual Reality baby nursery simulator that allowed users of all ages to experience what it was

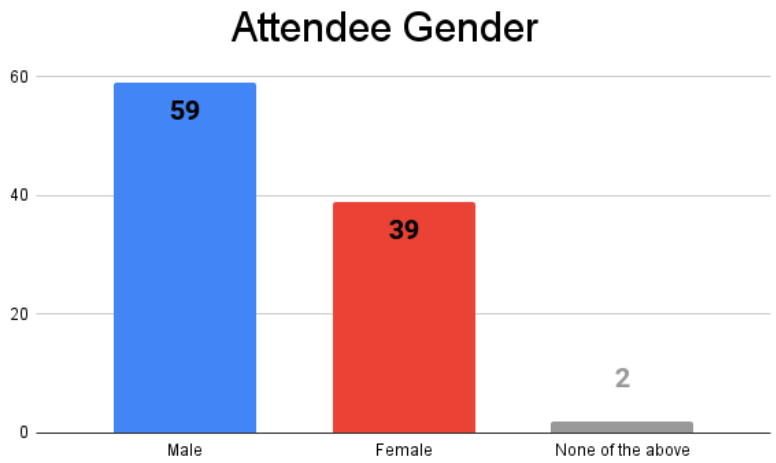

図 3.14　SIGGRAPH 2017 MQR *Real Baby - Real Family* Attendee Gender Survey

like to take care of a child. First, the users have a photo of their face taken which was then used to generate the face of the virtual baby, therefore having a virtual baby avatar resembling the user and enriching the experience. Throughout the play through, users get to interact with their virtual baby and engage in a series of interactions such as feeding the baby milk bottle and putting the baby to sleep.

Here, MQR was used to obtain both personal data as well as the type of baby they'd want to play, therefore having a more direct influence on the gameplay.

According to Figure 3.14, 59 out of 100 users were male while 39 were female and 2 identified as none of the above. This information was used to help better generate the facial image of virtual baby since users had their photos taken and then had their specific facial features extracted from these photos.

Since Real Baby project also allowed user to pick the type of personality and gender their baby possess, it was important to record such information before the experience. As we can see from Figure 3.14, the Real Baby project allowed for 11 types of personalities of the baby to be chosen, and the users answers were all documented in the MQR system. Here, the most widely chosen personality was "cheerful". We could also see from Figure 3.16 that 38 out of 100 users preferred a "girl" baby, 30 indicated "no idea", 29 preferred "boy", while 3 selected "others".

MQR was also used to record users' prior experience with infants and if they'd want

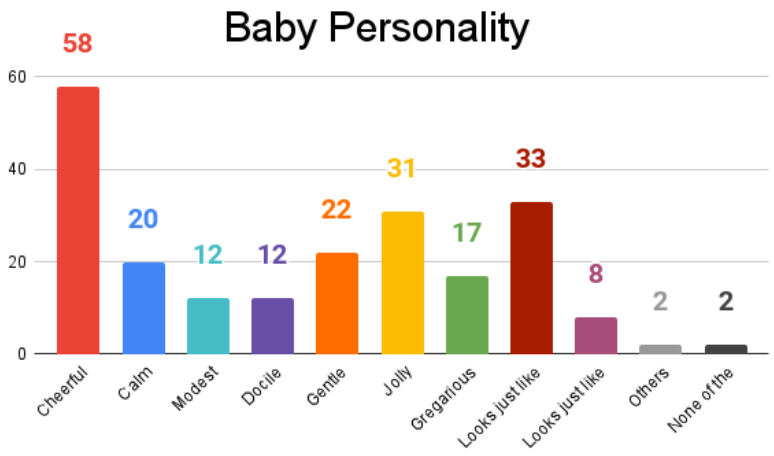

図 3.15　SIGGRAPH 2017 MQR *Real Baby - Real Family* Baby Personality Survey

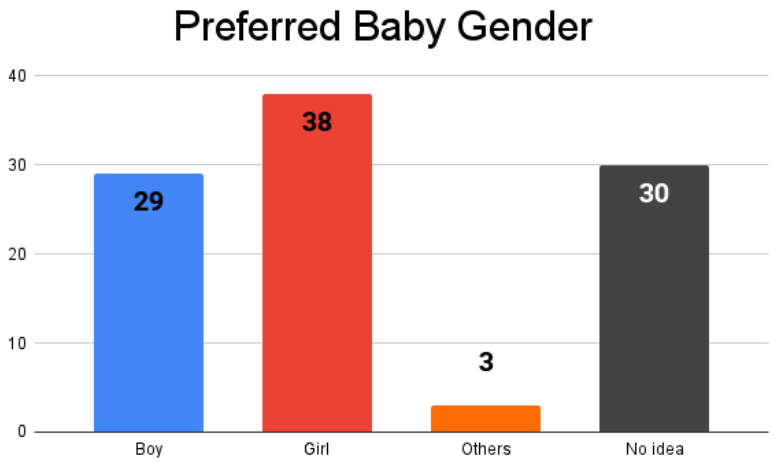

図 3.16　SIGGRAPH 2017 MQR *Real Baby - Real Family* Baby Preferred Baby Gender

to have their babies. According to Figure 3.17, 34 users indicated they have their children while 66 indicated they do not have their children. On the subject of "Do you want your own baby" shown in Figure 3.18, 54 indicated "Yes I do", 20 indicated "No I don't", another 20 indicated they wanted both imaginary baby and real baby, while 6 indicated they only wanted an imaginary baby.

As we can see, MQR can not only record users' information and evaluate their personal experience and thoughts regarding a virtual reality project, but it can also be used to record user data that has direct impact on the VR experience.

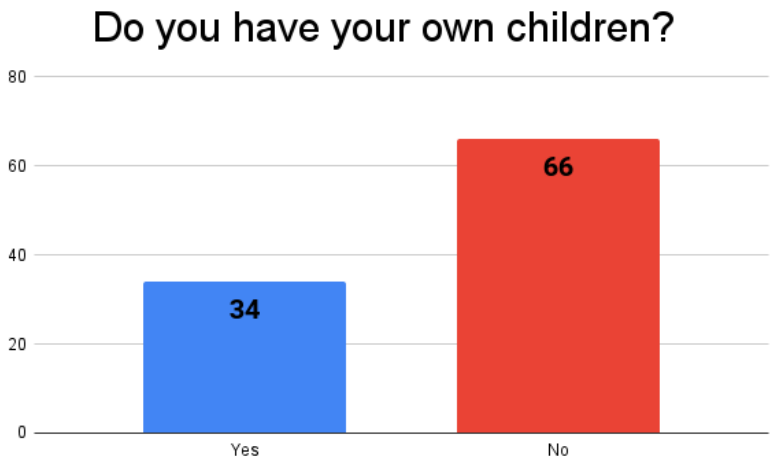

図 3.17　SIGGRAPH 2017 MQR *Real Baby - Real Family* Having Own Children?

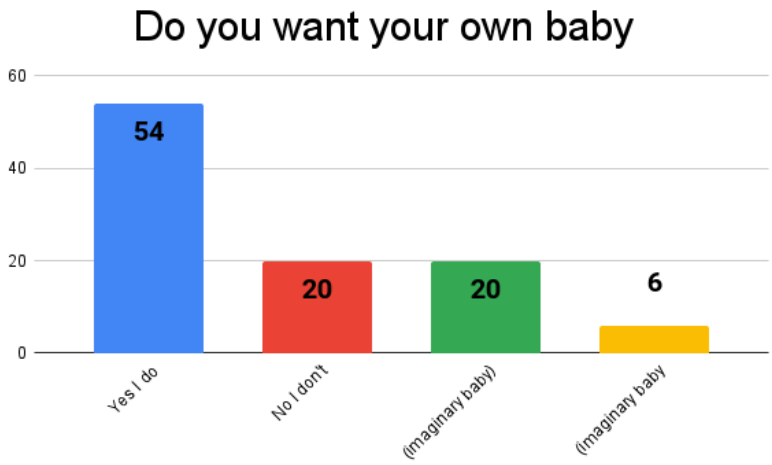

図 3.18　SIGGRAPH 2017 MQR *Real Baby - Real Family* Wanting Own Children?

3.6　MQR Contributions and Results

MasQueRade allows for the onsite evaluation of Virtual Reality systems and can even be used to record user choices that can alter user experience. It can be used in large scale exhibitions such as Anime Expo 2017 and SIGGRAPH 2017, recording over hundreds of user feedback that were later used to improve the VR systems. Furthermore, it can be used to record user inputs that altered the in-game experience of VR applications. Due to its scan and evaluate anytime nature of MQR, this onsite

evaluation system can be used to save time and manpower in large exhibitions where users otherwise have to wait in long lines to do the traditional paper-based evaluation.

第 4 章

eLearning anime Avatar video system

4.1 Introduction

This section describes an eLearning anime avatar video system that was introduced in 2019 with the purpose of assisting university students in learning in-class materials. In this research, videos featuring anime styled avatars were created and made available on class websites. Students were asked to take surveys both before and after viewing these videos to better access their knowledge retention.

The motivation behind this research is to try to introduce VTuber avatars, an immensely popular form of YouTuber that utilizes motion tracked digital anime avatars and sometimes machine altered vocals in place of real broadcaster visual and audio, into the eLearning environment. eLearning systems are wide spread in both the academia and industry and have its roots in Computer Assisted Instruction (CAI) [3], [33]. The researchers believed since VTuber is such a popular trend in youth culture, they will help boast academic performance of students.

4.2 Past Researches

Numerous researches have been done regarding how to best utilize eLearning materials and course structures. First instance of eLearning research was in 1978 [31] and since then, the field has gained substantial traction as technology advances. Zou et al., 2017 [4] created and evaluated what was called the Mulsemedia (Multi-Sensational Media) aimed at providing supplementary video material to aid the education of STEM course works. Several other researchers such as Giannoukos et al., 2008 [5], Kim et

al., 2009 [6], and He Ueno, 2012 [7] created complete online networks and attempted to promote more collaboration and sharing of contents. Intelligent tutoring systems is another form of eLearning that provides real-time feedback to students [46]. Amy Baylor [24] also suggested the use of agents able to generate desired emotion within students as a way to improve knowledge absorption. Video conferencing is another popular form of eLearning and has been used in dental schools [25], university setting during COVID19 shutdown [27] which is also known as Emergency-Remote Teaching (ERT) [29], between school teacher and teaching consultant [30], or in graduate school settings [26], [28]. Others have created a mix between gamification and Learning Management System [32] and used virtual interaction in addition to online quizzes and discussions to facilitate learning.

The advancement in eLearning system has also led to an increase in literature evaluating the effectiveness of various eLearning components. One such study was by Scotty D. Craig and Noah L. Schroeder which utilized 3 different audios: classic and modern voice engine produced voice and human voice to measure the effectiveness of each audio have on learning [8]. The study utilizes program generated audio; however, and does not measure the effectiveness of directly altered human voice on learning and how audio when paired with avatars of different genders can have different effect on learning which is what this research seeks to answer. Other researches such as one by Walsh [9] tried to teach health care using eLearning contents including video but did not include a variety of avatars and audios. While some past researches have argued what avatar styles and whether naturalism or stylization is better as a pedagogical avatar [10] thankfully this really is not an issue in Japan as the anime style is a widely recognizable style and accepted style to the Japanese audience. Furthermore, while some previous researches such as one by Amy Baylor et al. [11] found that avatars sharing the same gender with the viewers were able to yield more influence, based on the universal popularity of female anime characters, the research team has decided to pit students with avatars of different gender regardless of gender matching.

A considerable amount of researches have also been conducted regarding the importance of the presence of a visual avatar instead of just audio or texts. Amy Baylor et

al.'s various research on virtual avatar demonstrated giving the virtual agent an image, let it be still image or animated avatars, not only gives credibility to the agent but also helps them deliver messages more effectively [12], [13]. Rinat Rosenberg-Kima et al. also demonstrated having visual avatars contributed to significant positive motivational outcomes [45].

4.3 Spring 2019 Processing eLearning

For the Processing class in the Spring semester of 2019, the research team utilized both motion tracked visual avatar and transformed voice to create weekly eLearning videos. Furthermore, for this class, students were divided into groups with each group assigned to specific videos, therefore limiting the types of avatar and audio students were exposed to.

4.3.1 Research Method

186 students from a Processing class at the Information Science Department of Kanagawa Institute of Technology participated in this experiment that lasts for 11 weeks running from April 2019 to July 2019. Every Friday, students were asked to watch videos containing one of the 6 videos labeled: RO, AO, BO, RT, AT, BT which stands for: Real Visual (Male Lecture's visual) Original Audio, Avatar A (Male Style Avatar) Original Audio, Avatar B (Female Style Avatar) Original Audio, Real Visual Transformed Audio, Avatar A Transformed Audio, Avatar B Transformed Audio as shown in Figure 4.1. The lecturer's visual and original audio were captured using Open Broadcasting Software Studio (OBS Studio). Avatar A and B were created using "REALITY" application on iPhone, which is a popular VTuber broadcasting platform that allows users to create their anime styled 3D avatar with fast facial capture that runs on 30 to 75 fps. The transformed audios were generated using Roland VT-4 vocoder, a popular voice transforming device commonly used by VTubers to change the users' voice into one matching their virtual avatar. Voice transformation was done by adjusting the pitch, formant, balance, or reverb on the VT-4 vocoder. The 3 avatars and how they appeared in eLearning videos can be seem on the Figure 4.2. During the

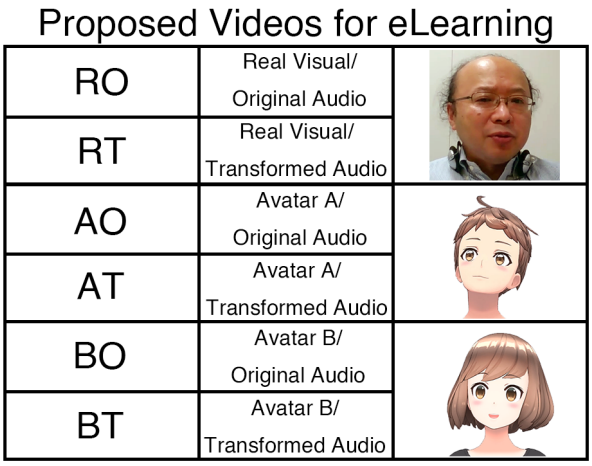

図 4.1　All 6 video types for 2019 Spring Processing Class

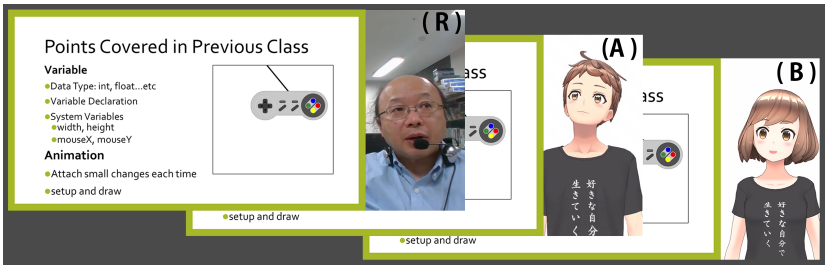

図 4.2　Video slides of 3 Avatars: (R)Real Avatar, (A)Male Styled Avatar, (B) Female Styled Avatar

video the avatar is on screen on the right side nonstop accompanied by lecture slides containing programming codes and concepts.

4.3.2　Student Division Method

The 186 students were divided into 15 groups labeled from A to O with each assigned to 1 or 2 videos as follows: [A: RO-RT, B: AO-AT, C: BO-BT, D: AO-BO, E: AT-BT, F: RO-AO, G: RO-BO, H: RT-AT, I: RT-BT, J: RO, K: AO, L: BO, M: RT, N: AT, O: BT]. Group A to I were asked to watch either of the 2 videos they were assigned to each week to measure their subjective impression differences for the 2 videos while groups J to O served as control group and watched only 1 of the 6 videos. The experimental groups A to I can be further classified into 3 main sections each designed to test a specific feature of the videos. Group A, B, and C, although watched different videos in odd and even week, did not have their visual's changed and the only difference

Groups	Odd Weeks (1, 3, 5, 7, 9, 11)	Even Weeks (2, 4, 6, 8, 10)	Purpose
A	RO	RT	Measure Audio with same visual
B	AO	AT	
C	BT	BO	
D	AO	BO	Measure Avatar Gender change
E	AT	BT	
F	RO	AO	Real Visual vs "REALITY" Generated Visual Avatars
G	BO	RO	
H	RT	AT	
I	RT	BT	
J	RO	RO	Control Group: Only Watched 1 Video
K	AO	AO	
L	BO	BO	
M	RT	RT	
N	AT	AT	
O	BT	BT	

図 4.3　Group division and videos assigned to each group

between the videos is their audio. Group D and E were designed to measure the difference between male and female styled avatar and thus the audio did not change in between weeks while the avatar switched between A and B. Group F, G, H, and I watched videos of the same audio but have their visual avatars switched between "REALITY" generated anime styled avatar and the real lecturer, therefore the visual change followed 1 of the 2 tracks: A - R - A, B - R - B. A detailed group division chart can be seem in Figure 4.3.

4.3.3 Subjective Data Gathering

Students attending the course were asked to fill out a survey before and after class as well as during each week after having watched the weekly eLearning videos in order to gather their subjective impression data. The surveys were structured in multiple choice (MCQ) or Likert Scale (LS) in a scale of 4 from "Disagree" (1) to "Agree" (4). The research team purposefully made the scale an even number to eliminate the neutral option. There was also a Short Answer Question (SAQ) inquiring about additional opinion. The survey questions for before class impressions are listed below.

1. Which video do you want to watch? (MCQ of 6)
2. I am looking forward to the class. (LS)
3. I have experience with VTuber. (LS)

4. I like Human Lecturer. (LS)
5. I know a lot about Processing. (LS)
6. I will take this class seriously. (LS)
7. I have experience with online lectures. (MCQ of 2)
8. Please give us any comments you have (SAQ)

The survey questions for the weekly videos are listed below. In order to make sure the students have finished each video and are viewing the correct video, we have asked each student to put down the start and end time as well as the keyword of the video they have watched.

1. Video start time (Time Input)
2. Keyword (SAQ)
3. How focused are you when watching the video? (MCQ of 6)
4. Video end time (Time Input)
5. The video's visual is good. (LS)
6. The video's audio is good. (LS)
7. The visual and audio do not mix well. (LS)
8. Overall the video is good. (LS)
9. Regarding the Audio (MCQ of 2)
10. Regarding the Learning Content (MCQ of 2)
11. Regarding the Avatar (Checkbox of 13)
12. Please give us any comments you have (SAQ)

4.3.4 Hypothesis

Anime characters, particular female anime characters have long since enjoyed high popularity amongst youth in Japan. Furthermore, the most popular VTubers with the most subscription in YouTube such as: Kizuna Ai (2 million subscribers), Kaguya Luna (1 million subscribers), Mirai Akari (734,000 subscribers), and Nekomiya Hinata (533,000 subscribers) are all female avatars. Due to the popularity of female avatars and the virtual non-existing of male avatars, the research team predicted that videos

featuring female Avatar (Avatar B) will be better received and will allow students to outperform the other groups grade-wise followed by Avatar A and finally Original Visual. The research team also predicted voice transformer will work best with avatar visuals but not with lecturer visuals. This hypothesis is drawn due to the number of success cases of male YouTube stars who have successfully transformed their voice to match their female styled avatar. Therefore, the predicted academic performance from best to worst is illustrated as follows: BT, BO, AT, AO, RO, RT.

4.3.5 Before Class Subjective Impression Results

Out of 186 students who were enrolled in the class, 182 students responded to the Before Class Survey. 174/182 (95.6%) of the students answered that they are looking forward to the class (rated 3 or 4 LS scale) while 172/182 (94.5%) said they will take the class seriously, pointing towards a positive attitude in the group. The majority of students also indicated they have no prior knowledge when it comes to eLearning, Processing or VTuber with 142 (78%) saying they have no experience with online learning, 169 (93.3%) saying they do not know Processing as shown in Figure 4.4, and 168 (91.2%) saying they do not have experience with VTuber as indicated in Figure 4.5. There is no clear indication of which videos the students want to watch. BO received the highest vote at 49 (27.1%) followed closely by RO and BT each at 48 (26.5%). RT received 20 (11%) of votes while AO and AT each got 8 (4.4%) according to Figure 4.6. There are also no clear indication of opinion regarding if students like human lecturer with 123 students (67.6%) answering either 2 or 3 on the LS scale.

4.3.6 Weekly Video Subjective Impression

According to the devised subjective evaluation using Google Form survey, the team has found Avatar B to be the most favorably ranked avatar amongst all three avatars while Avatar A and R were ranked just slightly behind Avatar B. Audio wise; however, Avatar B performs considerably less well in comparison to Avatar A and B ranking number 1 in the Question: “The visual and audio do not match. (LS)” as well as having the lowest score in “The video’s audio is good. (LS)” Original Audio also outperforms

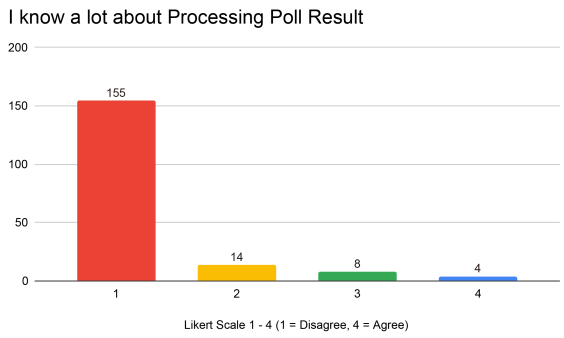

図 4.4　Before class processing knowledge opinion poll

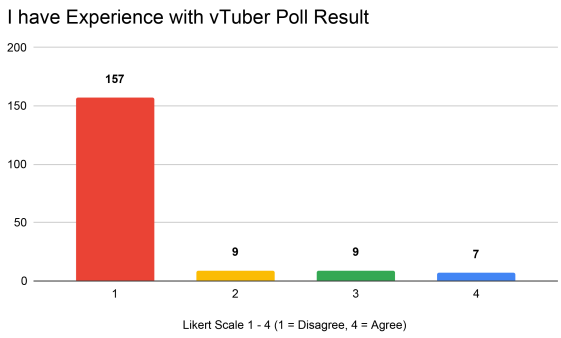

図 4.5　Before class VTuber knowledge poll

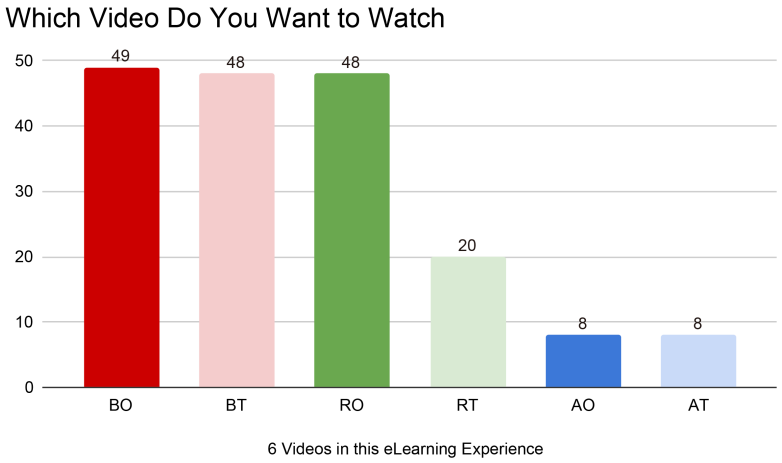

図 4.6　Before class “Which Video Do You Want to Watch” poll result

transformed audio by a significant margin and while O ranked 3 or above for “Audio is Good” Likert Scale Average, T only ranked around 2.5. The mismatch impression for visual and audio is also the best for RO at 1.85 while RT scored the 2nd at 2.25. BT was ranked as having the most mismatch at 2.90. Despite the difference in audio quality, all videos are rated a 3 or above in terms of impression for overall content. Figures 4.7, 4.8, 4.9, 4.10 are an average Likert Scale impression score of all 6 videos from students for the entire semester.

図 4.7　Visual impression Likert Scale score

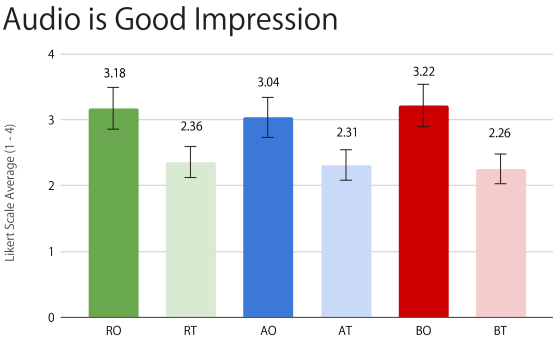

図 4.8　Audio impression Likert Scale score

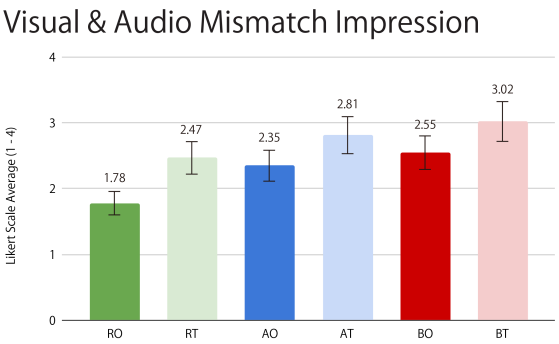

図 4.9　Visual and audio mismatch Likert Scale score

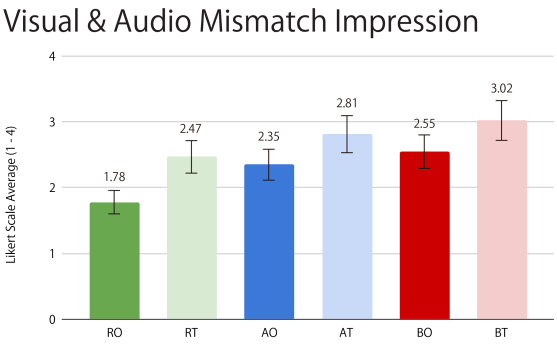

図 4.10　Overall impression Likert Scale score

4.3.7　YouTube Data Analysis

Aside from Google Form surveys, the research team also utilized YouTube Studio Analytics to conduct objective data analysis regarding the watch time and duration for the videos. The majority of the views took place on Monday and Thursday as shown in Figure 4.11, which were the days right before the processing class with Tuesday being the Processing Lecture and Friday being the Processing Workshop. Furthermore, by comparing the average view duration of each video with their length, the research team was able to obtain objective data regarding students' motivation towards watching each video. Overall, videos featuring original audio received longer viewing time in comparison to videos featuring transformed audio by an average of 16.87 seconds. While the video duration varies between each video, T videos, with the exception of A series have longer duration therefore taking away the possibility O receives longer views because they are longer when in fact O is shorter than T and yet receives longer

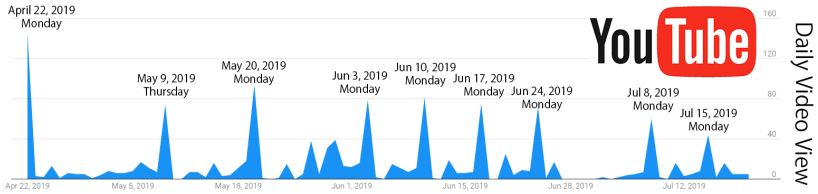

図 4.11 YouTube video watch rate from April to July 2019. Here it can be seem most video watches took place either on Monday or Thursday right before the lecturers.

views. This demonstrates audio quality is very important and can play a crucial factor in motivating students to watch them.

4.3.8 Objective Evaluation by Grade Performance

Students participated in this Processing Class were given 2 examinations and 1 project in the duration of 4 months with a mid-term exam on July 2nd, final exam at the end of July, and a final assignment that needed to be handed in at the beginning of August. This section will discuss the objective evaluation of these 15 groups using the averages of their grades for each assignment. This section also breaks down the 15 groups into their functions, for example groups A to C are the audio testing group where their visuals do not change and it is only the audio that changes from original (O) to transformed (T) in between weeks, and how groups of different functions fare in this eLearning environment.

4.3.9 Mid-term Grades

On July 2nd, 2019, a week between week 9 and 10 of the experiment, students were not assigned new videos or course materials due to them taking the mid-term exam. Of the 186 students originally registered for the class, 173 students completed the exam. This being said, due to some students being assigned into 2 groups, the grade average of each group and their ranking differs depend on whether these overlapped students' grades were used. The placing for each group with no overlapping student and overlapping student were shown in Figure 4.12 and 4.13 respectively. In both scenarios, E (BO/BT) ranked first, J (RO) ranked top 3, and I and O ranked bottom 3. Other groups' grades and placings changed dramatically depended on whether overlapped students were used

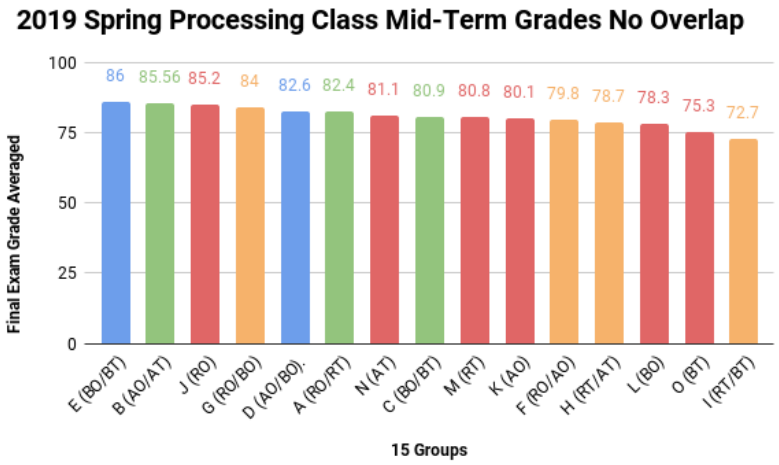

図 4.12　2019 Spring Processing class mid-term grades ordered by rank no overlap

or not. Most affected of these groups were C (BO/BT) going up from an average of 73 when including overlapped students to 80.9 when excluding 3 overlapped students. The average of 6 overlapped students was at 66.7 and will be excluded from further consideration due to them not solely belonging to any group in particular. This being said, according to T-Test, there's no statistical significance between E and C, the highest and lowest scoring group when taking overlap into consideration.

When taking only students who have participated in 1 group in consideration, the top performing measuring group were DE, the avatar gender change group with both placed top 5. The audio change group: ABC also performed well with all groups averaging above 80 and in the upper half. The control group and real avatar vs. digital avatar group all have mixed results with these 2 measuring group occupying the entire lower half of the 15 groups.

When comparing the various groups to control group (J – O), the research team found using T-Test that there were no statistical significance between audio test groups (ABC), avatar gender swap groups (DE), real avatar vs digital avatar groups (F – I), and control groups. There was also no statistical significance between control groups and the entire experimental groups (A – I). Furthermore, there was no statistical significance between E and I, the best and worst performing group just like when overlapped students were taken into consideration.

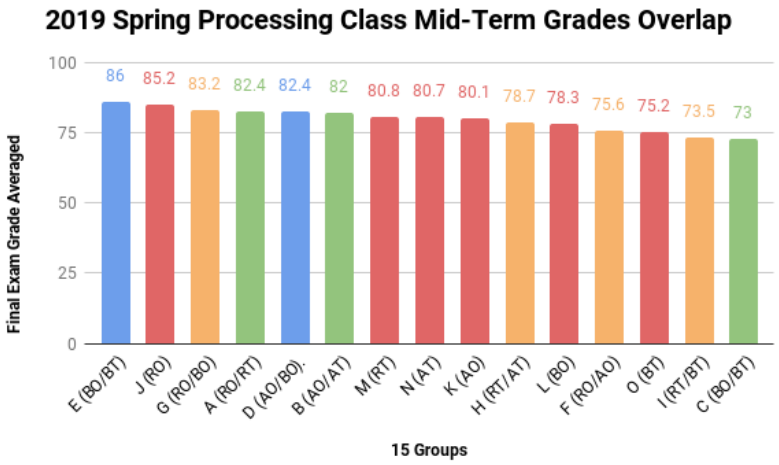

図 4.13　2019 Spring Processing class mid-term grades ordered by rank overlap

4.3.10　Final Exam Grades

The 2019 Spring processing class final exam graphs, both no overlap and overlap, can be seem in Figure 4.14 and 4.15. Here, unlike during mid-term where the ranking of different groups vary greatly depending on whether overlapped students were included, the different groups' final exam grade ranking stayed more or less the same with J (RO) ranked the highest and B (AO/AT) and I (RT/BT) ranked bottom 3. Between mid-term and final-exam, J and G were the two groups consistently placed highly with J placed top 3 and G placed top 4 in both cases.

The control groups (J - O) performed significantly better during final exam then they did during mid-term with all except O placed on the upper half. In regards to the different experimenting groups, the real avatar vs. digital avatar (F – I) also performed significantly better compared to during mid-term. The audio testing groups (ABC) would drop from all placed in the upper half to bottom 5 at the final exam. Avatar gender swap groups (DE) also experienced a drop in rank, albeit not as significant as the audio testing groups, from top 5 during mid-term to middle range.

A trend of students who only listened a single audio track performing better was also beginning to emerge during final exam as all upper half placed groups listened to one single audio track. It also became apparent that groups using transformed audio

(T) performed poorly with bottom 6 groups all used T either solely or interchangeably with original audio (O).

All of the groups assigned to measure audio consistency where their visuals do not change and only the audio changes are placed at the bottom half in terms of final exam grade performance. This trend is apparent in mid-term grades but not as apparent as A (RO/RT) did well at 4th back then. All groups performed not as well in final exam compared to mid-term with no exception. There is also T-Test statistical significance Control group and ABC (Audio Change group) while no statistical significance existed between Control group and other experiment groups.

In regards to visual avatar and grade performance, there didn't seemed to be a correlation between the two as no single avatar was placed consistently in groups performing well or unwell.

The students' final exam performance, in comparison to last year's performance, was overall better with the average grade increases by about 4 points from 65.6 to 69.6. The minimal grade has also went up from 5 to 24. It can be concluded that while different visuals and audios have different effects to students, the eLearning videos has been successful in helping students do better.

In regards to male vs female final exam grades, the average for the 145 male students who have completed the test was 70.12 with a standard deviation of 15.43 while the 20 female students have an average of 69.4 with a standard deviation of 12.07. There is no T-Test statistical difference between male and female students' final exam grades.

Amongst the 6 students who were assigned to more than 1 group (the overlapped students), the highest grade these students obtained was 78 while the lowest was 34 while the group average was 56.33. This, when compared to the average of 168 students who were only assigned to 1 group that was 70.3, was statistically significantly lower. This being said, due to the vast difference in number between group overlapped students and non-overlapped students, there existed the possibility had their numbers been more equal, the averages could have been closer.

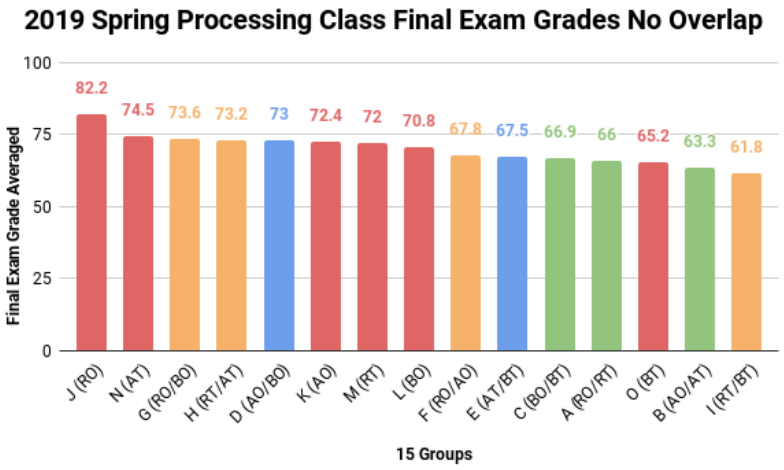

図 4.14　2019 Spring Processing class final exam ordered by rank no overlap

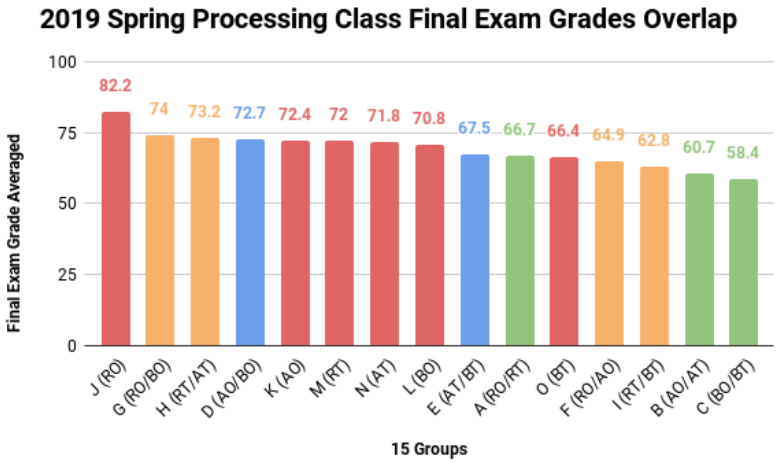

図 4.15　2019 Spring Processing class final exam ordered by rank overlap

4.3.11　Final Assignment Grades

The ranking of final assignment performance, both no overlap and overlap, can be seem in Figure 4.16 and 4.17 respectively. Here, the differences between overlap and no overlap grades were a lot smaller as aside from the bottom 5 groups, the rankings of all other groups stayed the same. Here, most of the control groups continued to perform well with M (RO), L (BO), and K (AO), J (RO) all received an average of 80 or above (J's 79.7 being the weakest one) while audio testing groups ranked poorly in no overlap

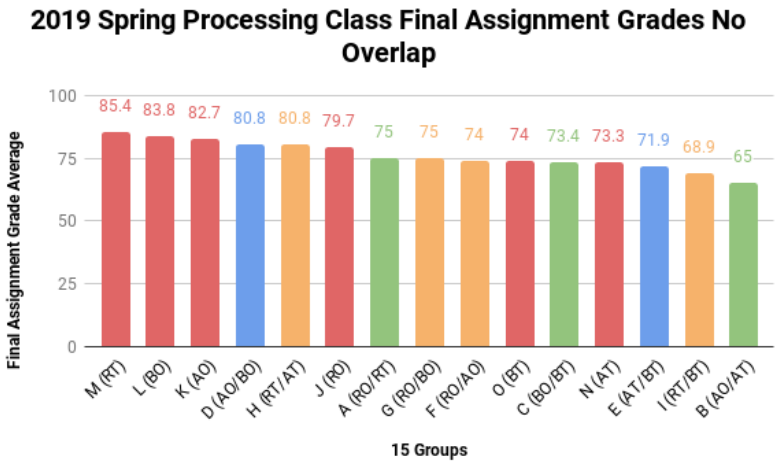

図 4.16　2019 Spring Processing class final assignment grades ordered by rank no overlap

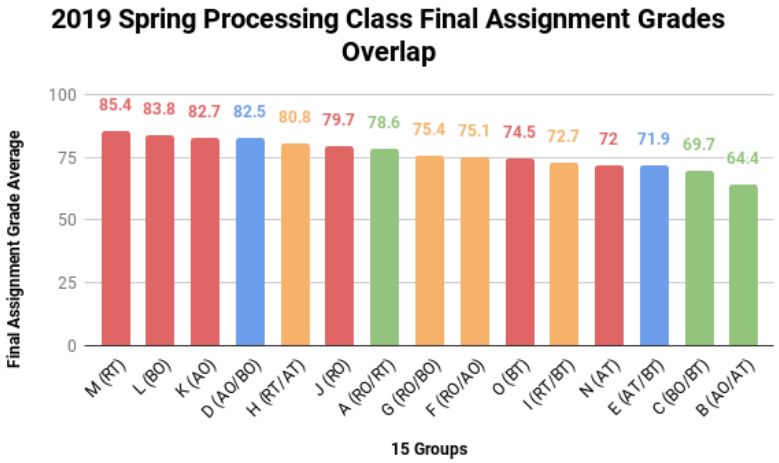

図 4.17　2019 Spring Processing class final assignment grades ordered by rank overlap

graph compared to other testing group with A(RO/RT) at 75 being the highest scoring one. It is also worth noting that the bottom 6 performers all had transformed audio (T) in them. This is not an indication that groups assigned to video with transformed audio (T) automatically performed poorly; however, as the highest scoring group M is assigned to RT alone despite the fact that the trend has been that controlled groups tend to do better grade wise. The other testing groups performed about average.

4.3.12 Final Grades

The ranking of final grades, both no overlap and overlap, are shown on Figure 4.18 and 4.19 respectively. According to the no overlap graph, 4 out of 6 control groups ranked at the better half and all 6 got an average of 70 or above. It can be concluded from the consistent performance of control group that students assigned to one single video featuring the same visual and audio performed above average in comparison to other groups.

It can also be seem that having consistent audio helped students to perform better grade wise as all audio experiment groups A (RO/RT), B (AO/AT), and C (BO/BT) were placed at the bottom 5 grade wise and all got below 70 in average with. According to T-Test, the research team also found audio experiment groups doing significantly less well compare to control groups while no other groups performed statistically significantly worse than control groups. Based on this finding, the research team can also conclude that inconsistency in audio affects anime styled avatars (A and B) more so that lecturer imageries as while group A (RO/RT) performed poorly this semester, sometimes it is ranked near the middle grade-wise while B and C composed of male and female styled avatar are placed consistently at bottom 2.

Groups assigned to the same audio (D - I) but different visual performed not as consistent as control group but not as poorly as audio testing group with I (RT/BT) placed bottom 3 while students from D to H all averaged within the 70s range. Based on these findings, it is safe to state audio impacts student performance much more so than visual with consistent audio being a very important factor in helping students learn effectively in an eLearning video environment. Original audio is also more beneficial to students' grade performance than transformed audio but the impact is not as strong compared to groups assigned to changing audio.

Despite of the inconsistency in grade performance, all of the groups have an average of above 60 meaning except for a few outliers, most of the students who participated in this eLearning experiment passed the course.

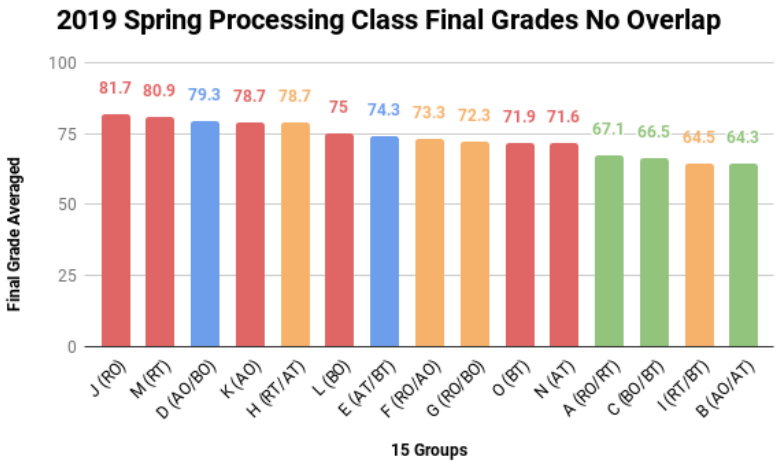

図 4.18　2019 Processing class final grades ordered by rank no overlap

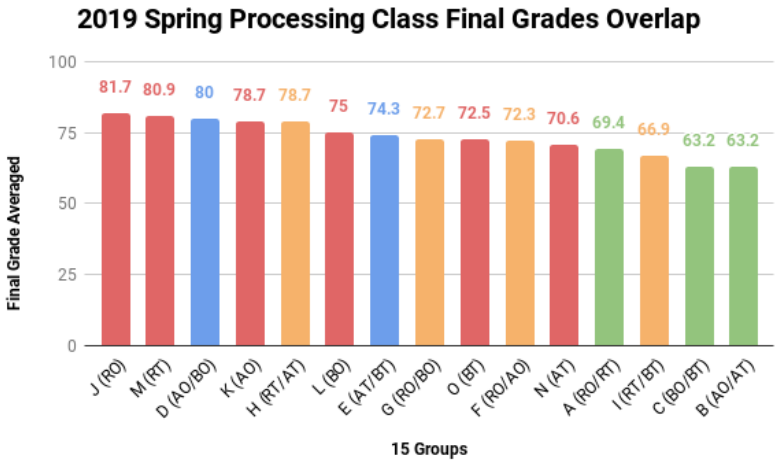

図 4.19　2019 Spring Processing class final grades ordered by rank overlap

4.3.13　Grade Comparison with Previous Semesters

In comparison to the grade performance of year 2015 to 2018, the grade of 2019 is ranked 2nd best in all categories, out-performed only by year 2016 except for the lowest grade received, which is the highest at 36 and 8 points higher than 2016, the year with 2nd highest minimum score. 2019 has the 3rd lowest SD value in the past 5 years losing to year 2015 and 2016. When comparing 2019 grades to only those of the past 3 years; however, 2019 performed the best in all categories. This shows a significant

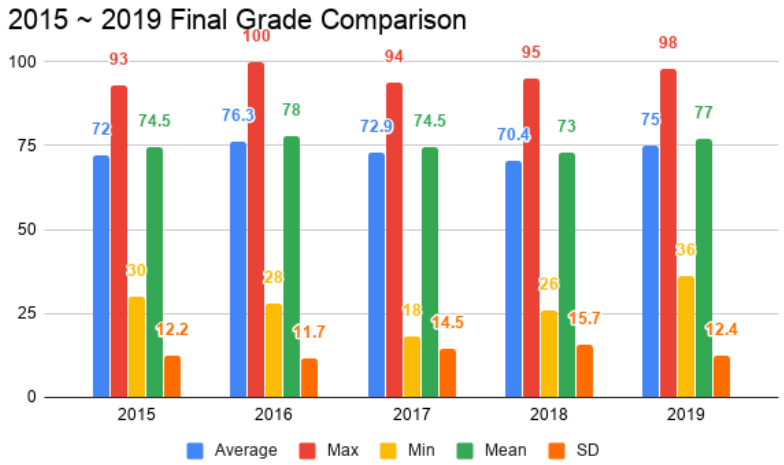

図 4.20　Processing class grade comparison within past 5 years (2015 – 2019)

improvement in grade performance thanks to eLearning videos. The data analysis here is based off of Figure 4.20.

4.3.14　eLearning video participation rate

Out of the 186 students in the class, 160 students participated in the 1st week of eLearning video experiment. By the 11th week, only 63 students remained. Of these 186 students, 20 of which (10%) were females while the rest 166 are males (90%). The overall participation number managed to remain at a healthy 100 or above until week 8 before dropping down to 91 for the first time in week 9. Participation took a drastic dip in the last 2 weeks to 75 in 10th week and 63 in 11th week. In the male population, the participation rate dropped from 140 (84.3%) to 52 (31.1%) while female participation rate went from 20 (100%) to 11 (55%). This showed that the experiment was more interesting to female population as motivation to watch videos for males quickly dwindled. This being said, we have to take into consideration that only 10% of the students were female and with such a small population pool when compared to the 160 male students, it is possible that with a more balanced male to female distribution, the participation of female can change. Data presented in this section is shown on Figure 4.21.

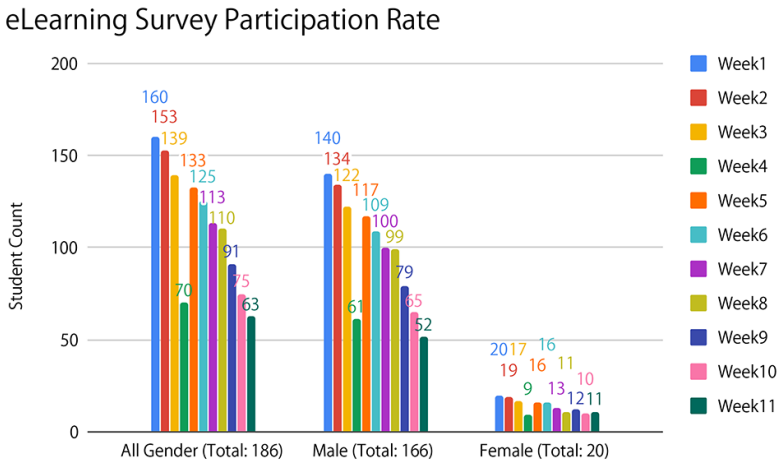

図 4.21　2019 Processing Class weekly student participation number

4.3.15　Hypotheses and Findings

Based on the findings of this experiment, the research team concluded that the 1st hypothesis regarding female avatar being the most well received avatar and one that helps with grade performance is only partially true in that while female styled avatars were rated as the most visually appealing avatar and with highest overall content score as well as longest video view duration out of all visual avatars, it did not help with grade performance as Real Avatar R coupled with constant weekly audio (regardless of the audio being O or T) were what helped students to perform well in class. The 2nd hypothesis regarding audio working the best with avatars is proven false as there is a higher amount of audio mismatch score between avatars paired with transformed voice than avatars with original voice. The data also pointed out that transformed audio overall scored lower points when compared with original audio without exception.

4.3.16　Limitation

Due to this research using academic grade to measure effectiveness of avatars, it must be noted that while grade improvement is a good indicator that visual avatars were able to aid students in learning, measurement based on grades in a course does have problems. Grades measured more than just the agent impact, and they are not

valid/reliable like a typical instrument would be in a study. While this study did utilize video watch duration and student impression in addition to grades, it did not have a clear pre/post video test directly measuring knowledge retention from each video.

4.3.17 Future Work

This research was conducted in a class composed of 90% of male students and 10% female students and therefore was not held in a gender neutral or even more female dominated population class so it awaits to be seem if female avatar will still be the most highly rated avatar visually and the best in terms of increasing video watch duration. Furthermore, the transformed audio was done by transforming the voice of a male lecturer in his 50s and it is highly possible for students to rate transformed audios differently had it been done by a female lecturer or a younger male lecturer. There were also no immediate ways of measuring knowledge retention in the current research as the mid-terms and finals all took place at least a week after video viewing. In the future, quizzes right before and after video viewing will be implemented to better test knowledge retention between avatars of different genders.

4.3.18 Conclusion

The proposed method can analyze interests and video viewing behavior of students when it came to preferred visual avatar, audio, and study time. This method contributed to the development of next generation of eLearning program utilizing motion tracking generated avatar visual and augmented audio by analyzing the most effective avatar and audio that both motivate students and enhance academic performance. Due to how both the visual and audio were generated in real-time by tracking the visual and audio outputs from the lecturer, this data analysis method has the potential to help develop live interactive education environment.

4.4 Fall 2019 POV eLearning

The POV-Ray eLearning videos were created with the shortcomings of previous experiment in mind. The usage of transformed audios were halted due to the potentially

adverse effects transformed audio has on students' grade performance. Furthermore, students were now free to choose the videos they wanted to watch and were imposed no limits on how many videos they could watch per week. Also in contrast to the previous semester where all weekly course contents were compiled into one single video, the different subject matters were now covered by different videos each with its specific theme.

The virtual avatars A and B were created using a popular 3D avatar tool named *VRoid*, a popular 3D model creation software in Japan and was then animated using *3tene*, which allows for the animation of 3D models using previously recorded motion tracking data. This ensured that all videos of the same theme but featuring different avatars were of the same length. This also decreased the work load of the lecturer considerably by allowing him to film for 1 video per subject and then using the facial data to create 3 separate videos.

4.4.1 Subjective Evaluation Survey

LS visual impressions score of the 3 avatars saw female styled avatar obtaining the highest score followed by Avatar A and R just like the Processing Programming Class in Spring semester. All 3 avatars received the score of 3 or above. This ranking is exactly the same for audio impression score with B >A >R which deviates from the trend in Spring semester; however, in that it was R that scored the 2^{nd} and A the 3^{rd}. Figure 4.22 represented LS visual score in Fall POV-Ray Tracing Class. Mismatch impression score in Fall semester shares the same trend as Spring semester with R (1.9) >A (2.38) >B (3.12). Furthermore, the differences between each avatar is much more pronounced compare to visual and audio score where all 3 avatars are graded at a different grade range. Overall impression score of the 3 avatars is R (3.44) >A (3.23) >R (3.17) which fits the trend of Spring semester counterparts. All 3 avatars are rated a 3 or above. The LS focus score of students is a new category introduced in Fall semester. Here the score ranking is as follows: B (3.44) >A (3.08) >R (3.02). While all 3 avatars once again received a score of 3 or above.

In summary the LS impression scores in Fall POV-Ray Tracing Class across all cate-

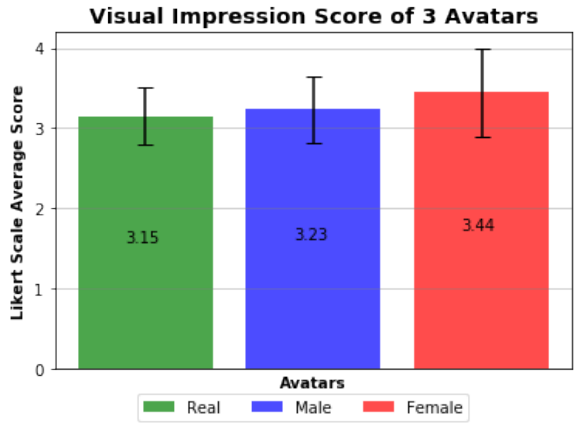

図 4.22 LS Visual Score of 3 Avatars in Fall POV-Ray

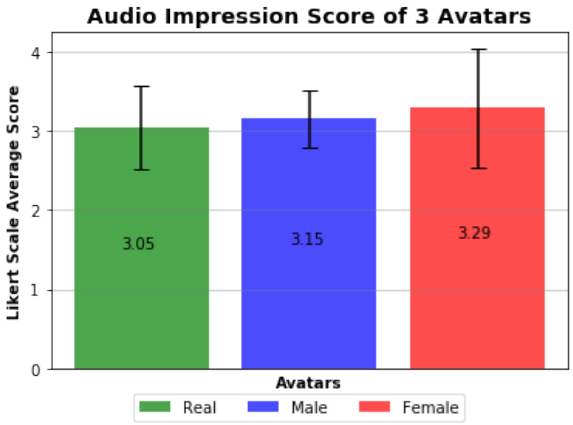

図 4.23 LS Audio Score of 3 Avatars in Fall POV-Ray

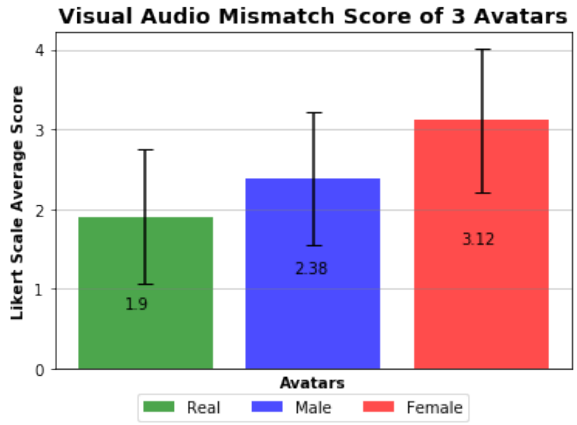

図 4.24 LS Mismatch Score of 3 Avatars in Fall POV-Ray

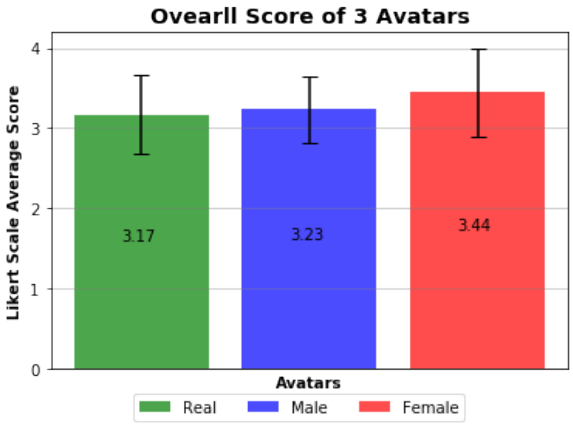

図 4.25 LS Overall Score of 3 Avatars in Fall POV-Ray

gories match that of Spring Processing Programming Class with the exception of audio score. In the newly introduced focus level survey, B performed the best followed by A and R. The result came as both a surprise and non-surprise due to how B generally performs better than other avatars in all other categories but the research team was thinking anime avatars may serve as distraction for students which was proven to be not the case. The data for this section can be seem in Figure 4.23, 4.24, 4.25, 4.26.

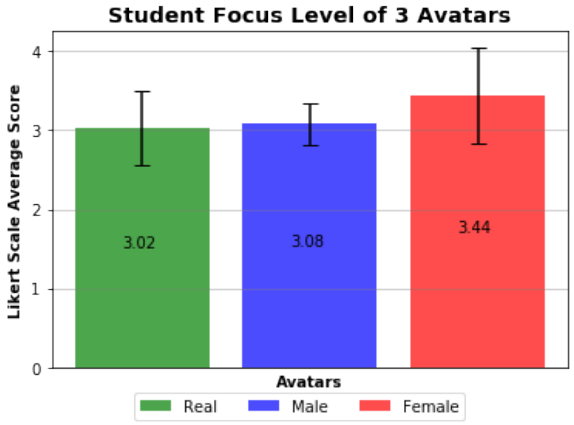

図 4.26　LS Focus Score of 3 Avatars in Fall POV-Ray

4.4.2　Fisher's Test: Good vs. Bad

To further test how the student population perceived each avatar in the visual, audio, mismatch, overall, and focus categories; the research team has decided to conduct Fisher's Test on all Likert Scale scores. The goal is to determine whether students consider the avatars good or bad and to achieve this, impression scores from 1 to 2 are considered bad while 3 to 4 are considered good.

This is the Fisher's Test score for visual impressions. For Avatar R, all 100% of both male and female students said the visual is good. For Avatar A, all 100% of male students said the visual is good while either no female students watched or gave Avatar A a visual rating. For Avatar B, 87.5% of male students said the visual is good while 100% of female students said visual is good. The Fisher's test visual data can be seem in Figure 4.27, 4.28, 4.29.

This is the Fisher's Test score for audio impressions. For Avatar R, 90.9% of male students said audio is good while 100% of female students said audio is good. For Avatar A, 100% of male students said the audio is good while either no female students watched or gave Avatar A a visual rating. For Avatar B, 75% of male students said the audio is good. 86.96% of female students said audio is good. The Fisher's test audio data can be seem in Figure 4.30, 4.31, 4.32.

This is the Fisher's Test score for mismatch impressions. Due to this being the

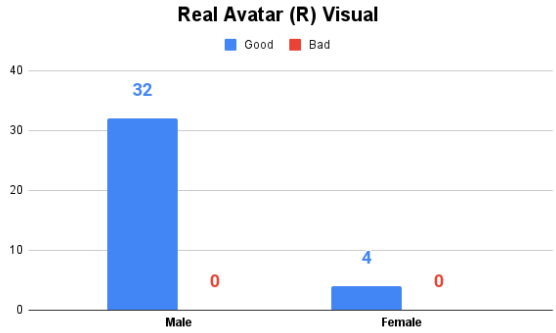

図 4.27　Fisher's Test Score of Avatar R's Visual impression divided by student gender

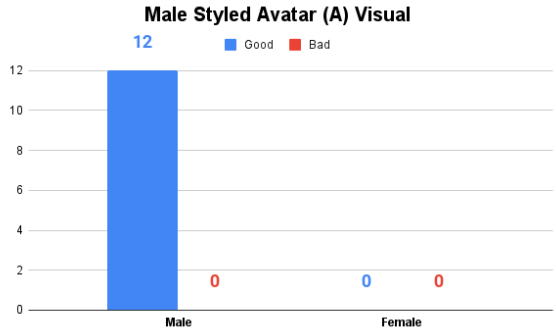

図 4.28　Fisher's Test Score of Avatar A's Visual impression divided by student gender

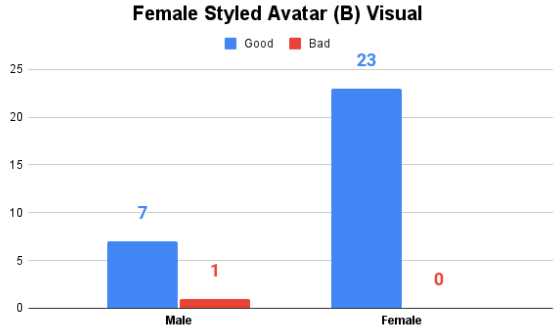

図 4.29　Fisher's Test Score of Avatar B's Visual impression divided by student gender

mismatch score, lower ratings of 1 and 2 means the mismatch is good while higher ratings of 3 and 4 means the mismatch is bad. For Avatar R, 66.67% (22/33) of male students said mismatch is good. 100% of female students said mismatch is good. For Avatar A, 58.33% (7/12) of male students said the mismatch is good while either no

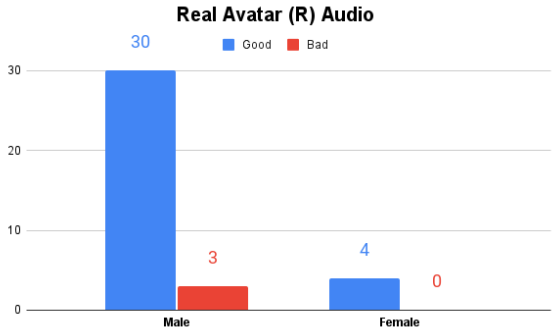

図 4.30　Fisher's Test Score of Avatar R's Audio impression divided by student gender

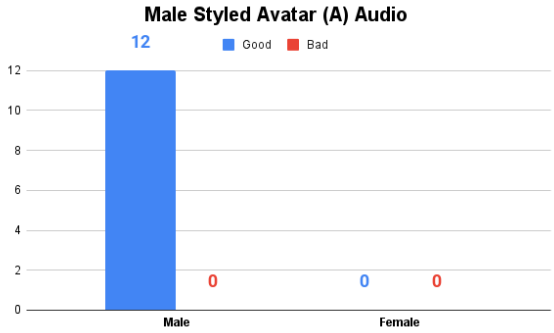

図 4.31　Fisher's Test Score of Avatar A's Audio impression divided by student gender

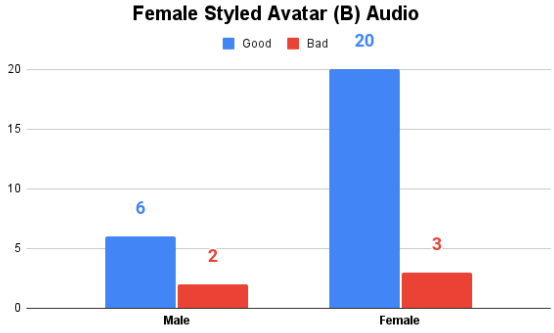

図 4.32　Fisher's Test Score of Avatar B's Audio impression divided by student gender

female students watched or gave Avatar A a visual rating. For Avatar B, 37.5% (3/8) of male students said the mismatch is good. 26.09% (6/23) of female students said mismatch is good. The Fisher's test mismatch data can be seem in Figure 4.33, 4.34, 4.35.

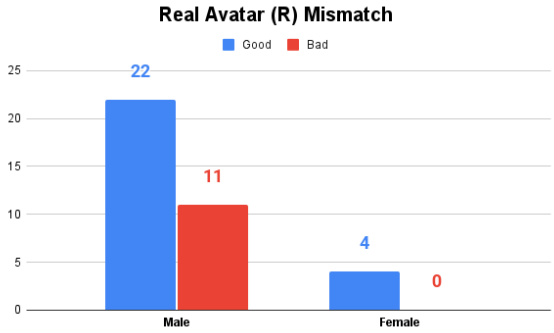

図 4.33　Fisher's Test Score of Avatar R's Mismatch impression divided by student gender

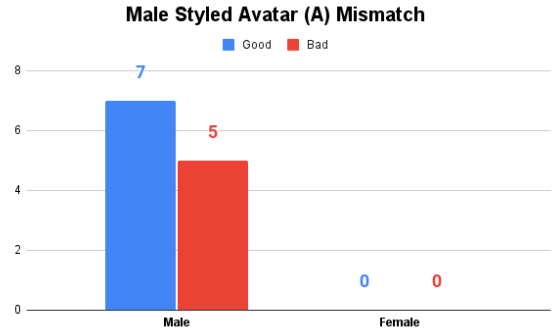

図 4.34　Fisher's Test Score of Avatar A's Mismatch impression divided by student gender

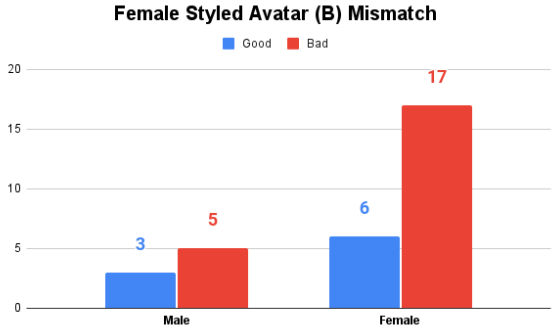

図 4.35　Fisher's Test Score of Avatar B's Mismatch impression divided by student gender

This is the Fisher's Test score for concentration impressions. For Avatar R, 90.9% (30/33) of male students said concentration is good while 100% of female students said concentration is good. For Avatar A, 100% of male students said the concentration is good while either no female students watched or gave Avatar A a visual rating. For

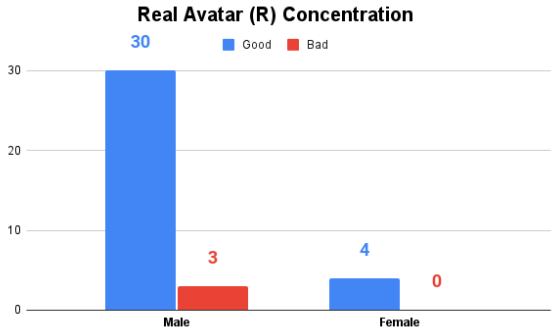

図 4.36　Fisher's Test Score of Avatar R's Concentration impression divided by student gender

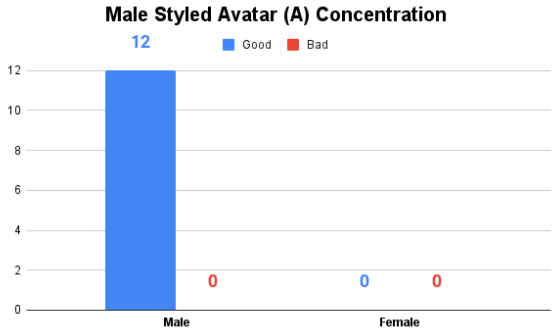

図 4.37　Fisher's Test Score of Avatar A's Concentration impression divided by student gender

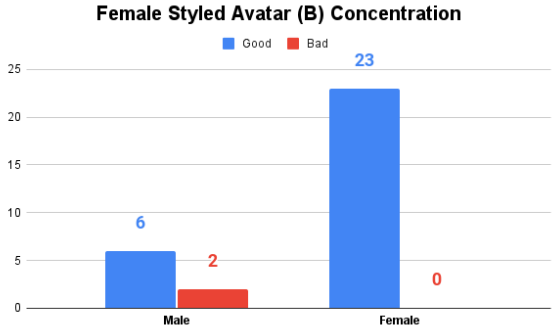

図 4.38　Fisher's Test Score of Avatar B's Concentration impression divided by student gender

Avatar B, 75% (6/8) of male students said the concentration is good while 100% of female students said concentration is good. The Fisher's test concentration data can be seem in Figure 4.36, 4.37, 4.38.

This is the Fisher's Test score for overall impressions. Avatar R: 96.88% (31/32)

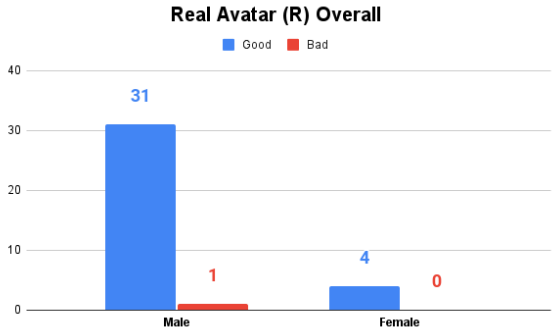

図 4.39　Fisher's Test Score of Avatar R's Concentration impression divided by student gender

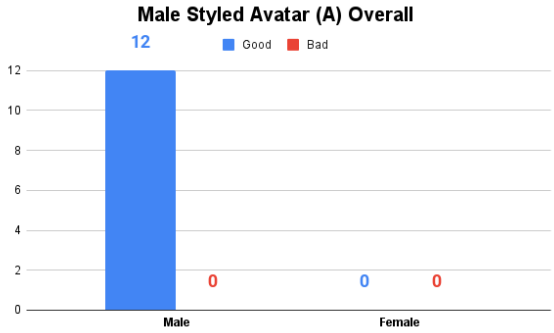

図 4.40　Fisher's Test Score of Avatar A's Concentration impression divided by student gender

of male students said overall is good. 100% of female students said overall is good. Avatar A: 100% of male students said the overall is good while either no female students watched or gave Avatar A a visual rating. Avatar B: 87.5% (7/8) of male students said the overall is good. 100% of female students said overall is good. The Fisher's test overall data can be seem in Figure 4.39, 4.40, 4.41.

Aside from the division of good and bad score between genders, another note worthy facet the research team would like to address is how different genders seem to prefer avatars of their own gender. While the Real Avatar, due to it being original visual of the male professor, was able to obtain views from both genders, the view rate is still disproportionately male at 89.19% (33/37). For male styled avatar (A), either no female students watched it or none who did rated the avatar. For female styled avatar, the research team predicted it would be the most popular of all avatars due to how well it performed in Spring 2019 Introduction to Processing class. Here, while it received

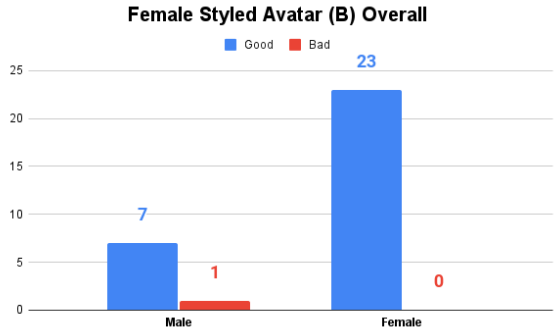

図 4.41　Fisher's Test Score of Avatar B's Concentration impression divided by student gender

more views from the same female gender at 74.19% (23/31).

This result seemed to reinforce Amy Baylor's conclusion that students are more drawn toward avatars of their own gender. To further test the viewing frequency of different avatars and even gender neutral avatars is something the research team would like to conduct in the future.

4.4.3　Fall 2019 Academic Performance

This section describes the academic performance of students in Fall 2019 POV-Ray Tracing class. Due to the nature of this class and how students are free to choose how many times they wish to watch the eLearning videos or not at all, the research team attempts to answer the question of whether the eLearning videos aided with better academic performance and if repeat watching of videos contributes to better grades.

4.4.4　Participants vs. Non Participants Grades

The grade average of all 198 students is at 72.61 and when subtracting that of dropouts is 76.88. The 23 participants of this eLearning program is at 89.22. The 175 non-participants' grade average is at 70.42 and within the 175 non-participants, the 164 non-dropouts' average is 75.15. According to T-Test, there is significant difference in grades between those who watched videos and those who did not with the exception of students who watched Avatar A videos. While this data seemingly shows that watching eLearning videos aided with better grades, currently there are no ways to

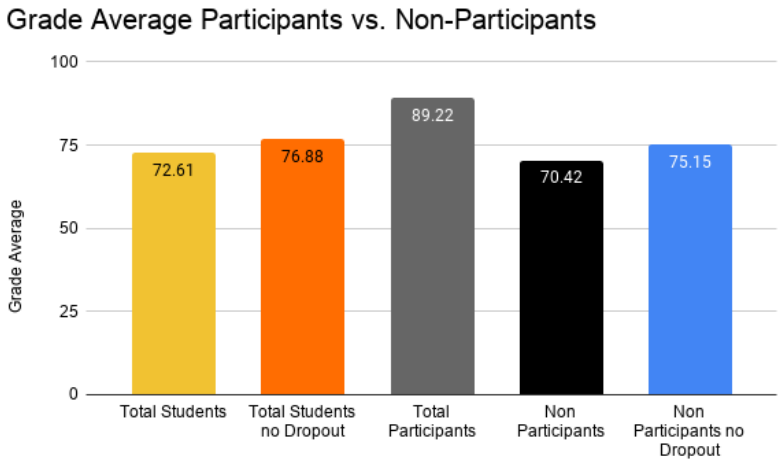

図 4.42　Grade average of participants

determine if naturally motivated students were more inclined towards watching eLearning videos or if the videos themselves helped students absorb knowledge. Figure 4.42 shows the grade average of different categories.

4.4.5　Watch Times and Grades

Of the 9 students who watched only 1 video in the entire semester, their grade average is 86.11. The 9 students who watched between 3 to 5 videos have the grade average of 89. The 5 students who watched more than 5 videos in Fall 2019 had the grade average of 95.2. According to T-Test, there is significant difference between those who only watched 1 video and those who watched more than 5 videos but not between any other watched times. Based on this finding, watching more videos does help with grade performance but a student has to watch significantly more than 1 video to see a meaningful increase in grade. All grade averages of participants were higher than the class average of 76.88. This data can be seem in Figure 4.43.

4.4.6　Avatars and Grades

The research team created 3 videos per subject matter each with their own avatar. This allows students to not only choose the area of study they want to engage in but to pick their own lecturers. Of the 23 participants, 9 only watched R, 8 only watched

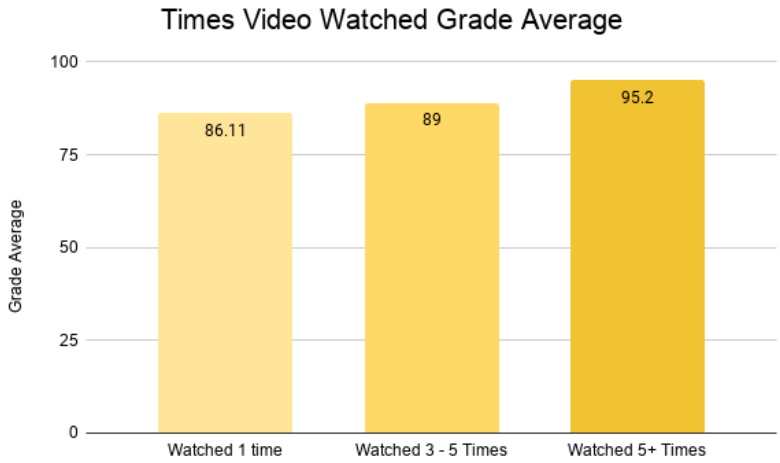

図 4.43　Video watched times and grade average

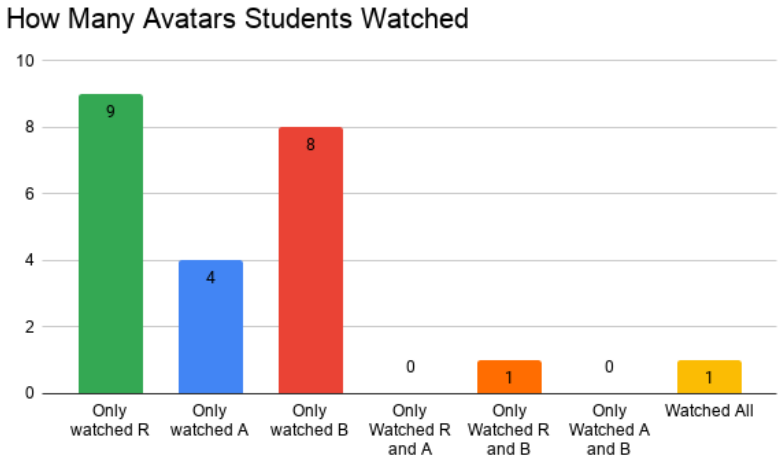

図 4.44　3 avatars watched times

B, and 4 only watched A, making R the most popular avatar followed by B and A. 1 student watched both R and B while 1 student watched all 3 avatars. This research not only shows the preference of avatars for students but also that most sticks with 1 single avatar and did not seem to enjoy variety in avatars as seem in Figure 4.44.

The grade average of students who watched each avatar is displayed on Figure 4.45 and their ranking is as follows: R (94.73) >B (88.4) >A (84). When the 2 students who watched more than 1 avatar is not counted, the grade average becomes: R (94.22) >B (86.25) >A (80) as indicated in Figure 4.46 and according to T-Test, there is statistical

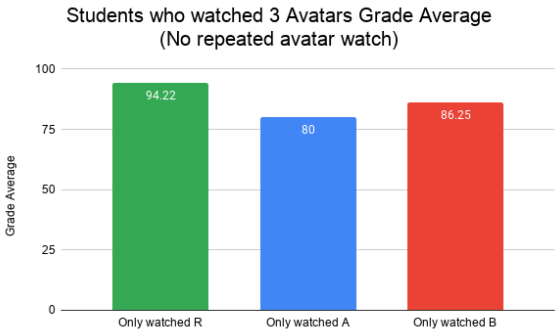

図 4.45　Grade Average of 3 avatars repeat viewers counted

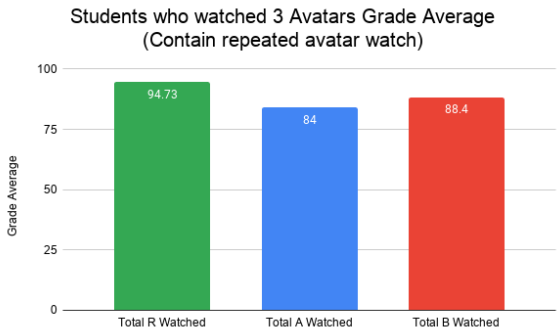

図 4.46　Grade Average of 3 avatars repeat viewers not counted

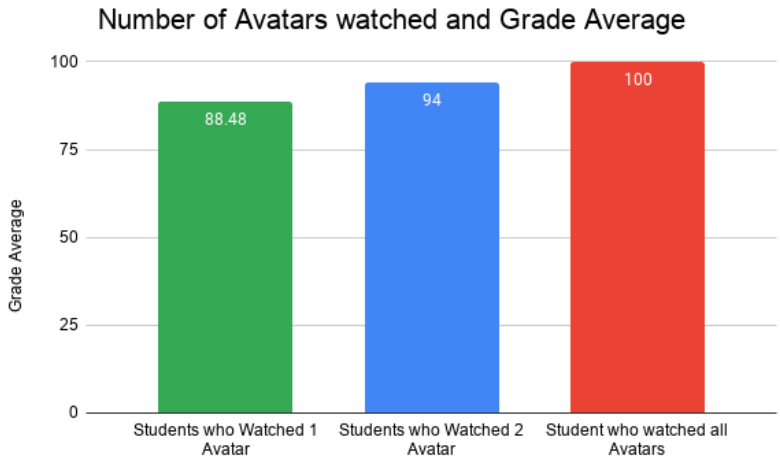

図 4.47　Grade average students based on variety of avatars watched

significance between these groups. The student who watched both R and B has the final grade of 94 and the student who watched all avatars has the final grade of 100 as seem in Figure 4.47. Due to only 2 students watched more than 1 avatar; however, the research team cannot conclude that watching a variety of avatars enhanced grade performance.

4.4.7　Gender and Grades

Of the 198 students in class, 173 were male and 25 were female, bringing the percentage of students of male and female to 87.37% and 12.63% respectively. 11 of the male students dropped out of class while none of the females dropped out, bringing

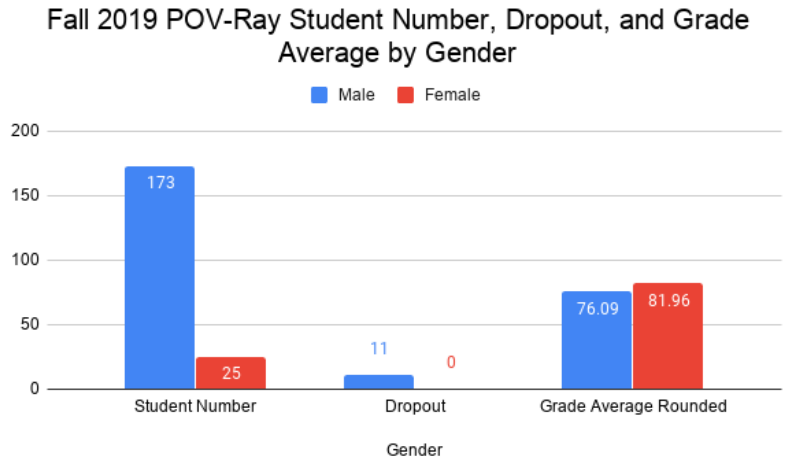

図 4.48　Grade average of male and female students

the total percentage of dropouts per gender to 6.36% for male and 0% for female. The grade average of male is 76.09 and female is 81.96. This data indicates that females outperform males grade-wise but it is not statistically significant and since the gender distribution of class heavily leans towards male, a better comparison would be if the male and female are distributed more evenly. Grade average of male and female can be seem in Figure 4.48.

4.4.8　Reasons for Picking Videos

Included in the survey is a section asking participants why they decided to pick the videos. The question is in MQC format and contains 4 choices: “No Specific Reason”, “Interest in Video Content”, “Like the Avatar”, and “I Simply Like it”. Of the participants who watched R, “No Specific Reason” received 20 votes, “Like the Avatar” received 12 votes, “Interest in Video Content” got 5 votes, and “No Specific Reason” got 1 vote. Avatar A got 10 votes for “No Specific Reason” and 2 for “Interest in Video Content”. B got 15 votes for “No Specific Reason”, 9 for both “Interest in Video Content” and “Like the Avatar”. Interestingly “No Specific Reason” received the most votes for every avatar closely followed by “Like the Avatar” and “Interest in Video Content” with the exception of A which received no votes for “Like the Avatar”. Figure 4.49, 4.50, 4.51 show watch reason for avatars R, A, and B respectively.

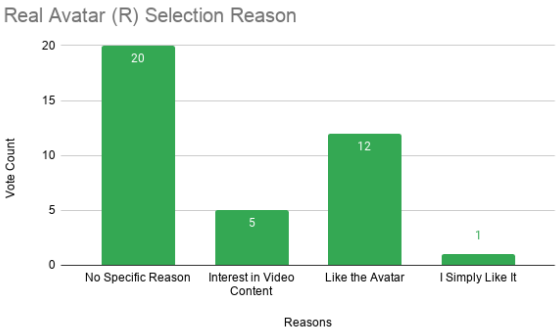

図 4.49 Real/Lecturer Avatar (Avatar R) Selection Reasons

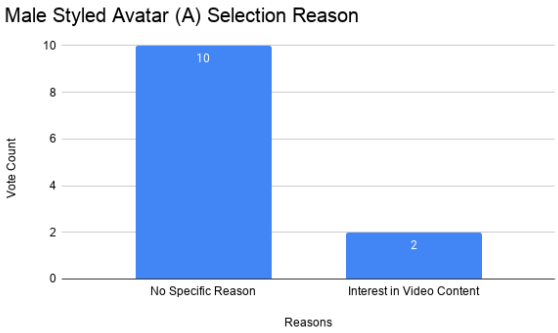

図 4.50 Male Styled Avatar (Avatar A) Selection Reasons

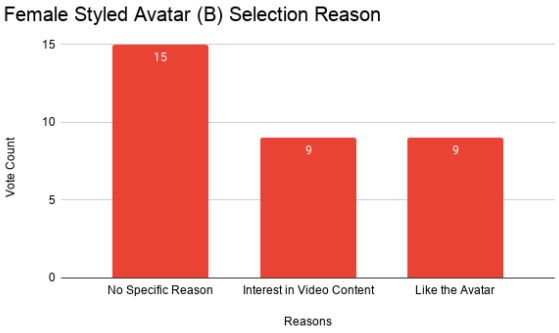

図 4.51 Female Styled Avatar (Avatar B) Selection Reasons

4.4.9 Fall 2019 Hypotheses and Findings

The research team predicted that Avatar B will be the most well received avatar followed by A and R. This hypothesis is proven correct according to the subjective evaluations where B is rated as the best character visual, audio, and overall wise followed by A and R. Only in visual and audio mismatch category was the ranking switched to R >A >B. Even in the focus level ranking it is B >A >R. This being said, when taken reasons for picking each video into consideration, 27.28% of students who watched B picked "Like the Avatar" while 31.58% of students who watched R picked "Like the Avatar", hinting at a higher percentage of students who watched videos featuring R due to them liking the avatar.

Regarding the hypothesis that R is best for students' grade performance, this was proven to be true as grade average of R watchers is at 94.22, topping A's 80 and B's

86.25.

For which avatar would be the most watched, it was found that R turned out to be the most watched avatar despite lesser subjective impression score given by students. It was also discovered that students seemed to prefer avatars of the same gender when given the chance to pick the video they would like to watch. While for the time being it would seem Amy Baylor's finding of students more comfortable with avatars of same gender, the research team does not rule out the possibility of all avatars voiced by the same male professor as reason for choosing videos and would like to explore visuals coupled with matching audios in the future.

As for the hypothesis: the more videos watched, the better their grades should get, this is proven to be true.

4.4.10 Conclusion

According to the 2 courses taught in 2019 utilizing avatar based eLearning videos, in a class that is predominantly male, Female Styled Avatar B was constantly ranked as the most appealing avatar with the exception of visual and audio mismatch. While the lecturer's original visual R is proven to be the best for grade performance and yields the most consistent results, Female Styled Avatars offer other benefits such as increase in watch duration and even focus level. When comparing grade performance with previous years as well as that of participants of eLearning programs with non-participants, it was found that the videos helped with increasing grade performance and the more videos students watched the better their grades tend to get.

In the future the research team planned on utilizing the eLearning videos in a more gender neutral class. Quizzes after each video view to measure knowledge retention will also be used to establish more direct relationship between video view and knowledge absorption.

第 5 章

Conclusion

This dissertation covered a new way to evaluate virtual reality applications that has proven to be useful especially at an exhibition setting along with an eLearning program utilizing VTuber styled motion tracked digital 3D avatars and transformed voices. This section will discuss contribution of the two key research areas in greater detail.

5.1 eLearning Program Empirical Contribution

The eLearning program contributed to the notion that avatars of specific styles and genders could encourage the prolonged viewing of educational videos. According to the research conducted in the year 2019, the use of anime avatars in Japanese education environment has yielded to roughly a 5% of increase in video watch duration. It was also found that transformed voice commonly utilized during VTuber broadcast in Japan did not yield positive results in neither students' academic performance nor video watch duration. Further, students in groups that watched videos with interchangeable audios do not perform as strong grade-wise as other groups with a unified audio. Due to the above findings, the research team was able to conclude that female styled avatar featuring original audio of lecturers was the best avatar to go about teaching contents in a University level eLearning environment but other forms of visuals including original lecturer's image and male styled avatar also aided in improved grade performance and watch duration. Transformed voice, regardless of whichever visual it was paired with, decreased watch duration but does not impact grade performance. This research proved that while visual is definitely important in encouraging the utilization of eLearning contents, a constant audio is the most important factor in grade performance regardless

of the style or modified method of audio.

5.2 MasQueRade Empirical Contribution

MasQueRade also contributed to a way of evaluating virtual reality systems by comparing user impression against after experience score, allowing the developers to get a clear sense of the users' thoughts on the product before and after experience. This is especially important for companies and developers who want to see if their pre-play demonstration and or advertisement resembles their content which, if there is too big a difference between actual gameplay and what is advertised, can lead to user dissatisfaction and a feeling of being lied to.

The portable nature of MQR and how it was linked to essential VR experience equipment also means all users can carry with them a system storing questionnaires for every single VR application in the event. This system can help event booths manage long waiting lines, save time for attendees who can enjoy other experiences after checking in by scanning the MQR system, and potentially help decrease paper usage by going from paper-based questionnaire to digital questionnaire.

5.3 Future Work

Despite of the positive results of anime styled VR avatar in a Japanese eLearning environment, more researches regarding the use of such avatar style in international stages await to be tested. Furthermore, studies regarding how international students in Japan approach anime styled avatars also awaited to be seem as it is unclear if children who may not have been exposed to Japanese anime style since childbirth would receive the anime avatars the same way.

Furthermore, due to the male dominate nature of the classes this eLearning system was utilized in, it remained to be seem if a more gender balanced environment will yield the same result.

While *MasQueRade* had been extensively tested in virtual reality exhibition and conference settings such as Laval Virtual in France and SIGGRAPH in USA, it is possible this remote evaluation system can be utilized in non-virtual reality settings.

It is possible to utilize this in a technology exhibition where name cards and even fliers that contain useful information and attendees would want to carry around with them during the entirety of the exhibition.

参考文献

[1] Aoto, T., Nakamura, Y., Harada, T., Ohkura, M. (2008). Study on Usage of Biological Signals to Evaluation Kansei of a System: Evaluation of feelings at a simple task using cards. Proceedings of the 2nd International Conference of Applied Human Factors and Ergonomics 2008.

[2] Ohkura, M., Aoki, Y., Aoto, T. (2009). Evaluation of Comfortable Spaces for Women using Virtual Environment -Objective Evaluation by Biological Signals-. Kansei Engineering International, 8, 67-72.

[3] Aparicio, M., Bacao, F. (2013). E-Learning Concept Trends. In Proceedings of the 2013 International Conference on Information Systems and Design of Communication (pp. 81–86). Association for Computing Machinery.

[4] Zou, L., Tal, I., Covaci, A., Ibarrola, E., Ghinea, G., Muntean, G.M. (2017). Can Multisensorial Media Improve Learner Experience?. In Proceedings of the 8th ACM on Multimedia Systems Conference (pp. 315–320). Association for Computing Machinery.

[5] Giannoukos, I., Lykourentzou, I., Mpardis, G., Nikolopoulos, V., Loumos, V., Kayafas, E. (2008). Collaborative E-Learning Environments Enhanced by Wiki Technologies. In Proceedings of the 1st International Conference on PErvasive Technologies Related to Assistive Environments. Association for Computing Machinery.

[6] Kim, J., Jung, Y., Lim, Y., Kim, M., Noh, S. (2009). An E-Learning Framework Supporting Personalization and Collaboration. In Proceedings of the 3rd International Conference on Ubiquitous Information Management and Communication (pp. 635–638). Association for Computing Machinery.

[7] He, Z., Ueno, H. (2012). Designing M-Learning System Based on WebELS Sys-

tem. In Proceedings of the 4th International Conference on Internet Multimedia Computing and Service (pp. 115–118). Association for Computing Machinery.

[8] Craig, S., Schroeder, N. (2017). Reconsidering the voice effect when learning from a virtual human. Computers Education, 114, 193-205.

[9] Kieran Walsh (2018). Reflections of health care professionals on e-learning resources for patient safety. Baylor University Medical Center Proceedings, 31(1), 35-36.

[10] Gulz, A., Haake, M. (2006). Design of animated pedagogical agents—A look at their look. International Journal of Human-Computer Studies, 64, 322-339.

[11] Baylor, Y. (2004). Pedagogical Agent Design: The Impact of Agent Realism, Gender, Ethnicity, and Instructional Role. In Intelligent Tutoring Systems (pp. 592–603). Springer Berlin Heidelberg.

[12] Baylor, A., Warren, D., Park, S., Shen, E., Perez, R. (2005). The Impact of Frustration-Mitigating Messages Delivered by an Interface Agent. In Proceedings of the 2005 Conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology (pp. 73–79). IOS Press.

[13] Amy Baylor, Jeeheon Ryu (2003). The API (Agent Persona Instrument) for Assessing Pedagogical Agent Persona . In Proceedings of EdMedia + Innovate Learning 2003 (pp. 448–451). Association for the Advancement of Computing in Education (AACE).

[14] 土佐 尚子 (1995). 音声で人間とコミュニケーションできるロボティクスアート: ニューロベイビー. 計測と制御, 34(4), 311-314.

[15] Metta, G., Sandini, G., Natale, L., Panerai, F. (2002). Sensorimotor Interaction in a Developing Robot. First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems.

[16] Ido, J., Ueda, E., Matsumoto, Y., Ogasawara, T. (2007). Robotic Telecommunication System Based on Facial Information Measurement. In Proceedings of the 12th International Conference on Intelligent User Interfaces (pp. 266–269). Association for Computing Machinery.

[17] Caregivers, H., Kozima, H., Yano, H. (2001). A Robot that Learns to Communicate with Human Caregivers. Proceedings of the First International Workshop on Epigenetic Robotics, 2001.

[18] Scassellati, B. (2002). Theory of Mind for a Humanoid Robot. Autonomous Robots, 12, 13-24.

[19] ONO, C., KUNIMURA, H., HIRAI, M., MATSUZAKI, W., MURAMOTO, M., UCHIYAMA, T., SHIRATORI, K., HOSHINO, J. (2009). P22 Design Consideration in Developing Baby Type Robot. Proceedings of the Annual Conference of JSSD(56), 354-355.

[20] Thies, J., Zollhöfer, M., Stamminger, M., Theobalt, C., Nießner, M. (2016). Demo of Face2Face: Real-Time Face Capture and Reenactment of RGB Videos. In ACM SIGGRAPH 2016 Emerging Technologies. Association for Computing Machinery.

[21] Shirai, A. (2016). Manga Generator, a Future of Interactive Manga Media: Invited Talk Paper. In Proceedings of the 1st International Workshop on CoMics ANalysis, Processing and Understanding. Association for Computing Machinery.

[22] Karen Simonyan and Andrew Zisserman (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. In 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings.

[23] Simo-Serra, E., Iizuka, S., Sasaki, K., Ishikawa, H. (2016). Learning to Simplify: Fully Convolutional Networks for Rough Sketch Cleanup. ACM Trans. Graph., 35(4).

[24] Baylor, Amy. (2018). Three Research Directions for Affective Learning Technologies.

[25] Oki, A., Sunariani, J., Yuliati, Y., Irmawati, A. (2020). Dental Students' Perception of Online Lecture Using Video Conferencing. Systematic Reviews in Pharmacy, 11, 245-248.

[26] Ghazal, S., Samsudin, Z., Aldowah, H. (2016). Students' perception of Synchronous courses using Skype-based video conferencing. Indian Journal of Science

and Technology, 8.

[27] Serhan, D. (2020). Transitioning from Face-to-Face to Remote Learning: Students' Attitudes and Perceptions of using Zoom during COVID-19 Pandemic. International Journal of Technology in Education and Science, 4, 335-342.

[28] Wang, Q., Huang, C., Quek, C. (2018). Students' perspectives on the design and implementation of a blended synchronous learning environment. Australasian Journal of Educational Technology, 34, 1-13.

[29] Nikou, S. (2021). Web-based videoconferencing for teaching online: Continuance intention to use in the post-COVID-19 period. Interaction Design and Architecture(s) , N. 47, 123-143.

[30] Fischer, A., Dart, E., Leblanc, H., Hartman, K., Olinger Steeves, R., Gresham, F. (2016). AN INVESTIGATION OF THE ACCEPTABILITY OF VIDEOCONFERENCING WITHIN A SCHOOL-BASED BEHAVIORAL CONSULTATION FRAMEWORK. Psychology in the Schools, 53, 240-252.

[31] Ball, G. H. (1978). Telegames Teach More Than You Think. Audiovisual Instruction. May, 2426.

[32] González González, C., Blanco Izquierdo, F. (2012). Designing social videogames for educational uses. Computers Education, 58, 250-262.

[33] Zinn, K. L. (2000). Computer-assisted learning and teaching. In Encyclopedia of Computer Science (pp 328–336). Chichester, UK: John Wiley and Sons Ltd. Retrieved from: http://dl.acm.org/citation.cfm?id=1074100.1074248

[34] Sutcliffe, A., Gault, B. (2004). Heuristic evaluation of virtual reality applications. Interacting with Computers, 16(4), 831-849.

[35] Bowman, D. (2002). A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. Presence: Teleoperators Virtual Environments, 11, 404-424.

[36] Mahatody, T., Sagar, M., Kolski, C. (2010). State of the Art on the Cognitive Walkthrough Method, Its Variants and Evolutions. Int. J. Hum. Comput. Interaction, 26, 741-785.

[37] Lewis, C., Polson, P., Wharton, C., Rieman, J. (1990). Testing a Walkthrough

Methodology for Theory-Based Design of Walk-up-and-Use Interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 235–242). Association for Computing Machinery.

[38] Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed.), Perspectives of curriculum evaluation, American Educational Research Association Monograph. Chicago, Rand McNally.

[39] Hix, D. and Hartson, H. R. (1993). Developing User Interfaces: Ensuring Usability through Product Process. New York, John Wiley and Sons.

[40] Nielsen, J. (1994). Usability Inspection Methods. In Conference Companion on Human Factors in Computing Systems (pp. 413–414). Association for Computing Machinery.

[41] Docebo (2016). ELEARNING MARKET TRENDS AND FORECAST 2017-2021. https://www.docebo.com/resource/elearning-market-trends-and-forecast-2017-2021/

[42] Japan Environmental Sanitation Center. (2014). History and Current State of Waste Management in Japan. https://www.env.go.jp/en/recycle/smcs/attach/hcswm.pdf

[43] Tortor, R., Coloma, B., Dizon, A., Granada, C., Lechuga, M. (2019). SIMULINE: Cloud-Based Waiting Line Management Tool with Mobile Technology.

[44] United States Environmental Protection Agency (2020). Paper and Paperboard: Material-Specific Data. https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/paper-and-paperboard-material-specific-data

[45] Rosenberg-Kima, R., Baylor, A., Plant, E., Doerr, C. (2007). The Importance of Interface Agent Visual Presence: Voice Alone is Less Effective in Impacting Young Women's Attitudes toward Engineering. In Proceedings of the 2nd International Conference on Persuasive Technology (pp. 214–222). Springer-Verlag.

[46] Ma, W., Adesope, O. O., Nesbit, J. C., Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. Journal of Educational Psychology, 106(4), 901–918. https://doi.org/10.1037/a0037123