Cybersickness Assessment on Virtual Reality

Deyuan Yang, Johnny Kim, Victoria Interrante

A B S T R A C T

It is commonly known that Virtual Reality (VR) is an important topic now in the area of technology. Although many people are keen on using VR device, they often feel uncomfortable with the VR when they see the moving scene. This research is to investigate how the VR affect people in cybersickness and what way we can use to reflect the degree of cybersickness correctly. An experiment set a basic environment for people in Virtual Reality and give them some tasks to measure their correctness of expression the brightness of sky in Virtual Reality environment by three different devices.

1. Introduction

Today, the technology of Virtual Reality (VR) becomes more and more popular in a variety of fields. People also have plenty of interests to explore more about the Virtual Reality. As the development of it, many problems appears while people using it, the symptom of cybersickness is one of which. When playing the VR game or watching VR movie, most of people feel sick, nausea and even vomiting, which are all caused by the scene of VR and the inevitable movement. So far, many research investigate the reason of causing the cybersickness and how to measure and deal with it. One of the experiments was to investigate the usage of a physical dial as a measurement tool for cybersickness measurement in Virtual Reality (McHugh, Hoermann, Jung, Lindeman, 2019). This article indicates that physical dial is a good way to convey the degree of cybersickness and also it mentions many other ways to test the cybersickness. The current gold standard metric for cybersickness assessment is subjective reporting. The two most commonly used instruments are the Kennedy-Lane SSQ (administered after an immersive experience) and the FMS (fast motion sickness) scale, administered at repeated discrete intervals during an immersive experience, such as every 2 minutes. We seek to learn patterns of data associated with impending cybersickness, to be able to predict cybersickness onset before it occurs. These efforts would be facilitated if we had access to continuous data about cybersickness outcomes along with continuous data about the visual stimulus (e.g. 3D optical flow, considering contrast and brightness) and the user's head/eye behaviors (e.g. instantaneous 6DOF head pose, eye gaze, and gradients thereof) as well as various physiological measures. This research help to explore a wide range of device and ways to test the cybersickness and check the validity of the result. In this experiment, we use button, slider and also knob as the tools to test the accuracy of the expression of participants. Before running people in the experiment, we set up a scene with sky and landscape with skull of goat in Unreal Engine. There are two tasks for each participant, one is to express the brightness of the sky and the other one is to pick up the goat skull as much as possible.

Due to the time limited, I only run 1 people as the participants to test the three devices and later the other PHD student will run more participants. By analysing the data, the three devices are all good to be used to test, but they have different benefit and deficiency.

2. Theory

During the experiment, each participate should sit on the seat and put on the Vive headset. This experiment requires users complete a search task within a virtual environment while reporting a number between 0-20 that corresponds to the current state of the changing weather conditions. The goal of this experiment is to compare the relative usability and cognitive overhead demanded by the use of three alternative devices for

* *Corresponding author.* E-mail address: auther@depauw.edu

reporting information using the non-dominant hand. We provide a training part for them to be familiar with the environment and their task and then in the formal experiment part, we will collect the 3 different data, including the score they gain for collecting skull part, position data which is the position they move and also their prediction to the brightness. By calculating the difference of the predication brightness and the actual data, we can conclude the most accurate tool for them to express the degree of cybersickness. Also, by testing the score data and position data, we could know the usage of the three tools.

After the experiment, we also make a small test to the participants to test whether they can find the number they want to express accurately and quickly by those three tools. This is also important to make us know the error and difference of the experiment.

3. Procedure, Results, and Discussion

Before the experiment, we should set up all the machine and vive. After the participant prepared, we should test their visual acuity first and to decide whether he or she meets the criteria of the experiment. And then, we have participant put on HTC Vive and calibrate IPD. After they fill out the consent form, demographics form and SSQ questionnaire, they can begin the training part of the experiment.

During the training part, we show them the brightness of the sky by 0, 5, 10, 15, 20 and let them have the sense of the brightness. Also, we tutor them how to use the trigger to capture the skull and also how to use the three devices one by one. After they feel like they are familiar with the environment and also the tools, we will start the experiment part.

In the experiment, we ask the participant to operate twice for each device. We won't give them prompt or reminder to change the number as the brightness changes, but just use their memory. After the main experiment, we also give the participant a small test for the three devices. According to the prompt on the screen, they will change the number of devices to get the number shown on the screen and compare the accuracy and speed of each device.

The data shows us a lot of information about the three devices. According to Figure 1 (a), the pattern of true answer and user answer is the same generally. The red line represents the true answer, and the trend of the red line and blue line is the same, which means that the brightness estimation of using the button is comparably accurate. During about 1000 milliseconds, the true brightness of the sky is about 10, but the participant chooses 3. The difference between the real and the user's idea is large. This means the recognozation of the user for brightness is not so good. This also can be shown in the later part of the experiment. Also, there is a possibility that the user forgot to change the brightness while they are doing the collecting skull task. According to Figure 1 (b), the pattern of the user answer is corresponding to the true answer generally. The red line represents the true answer, and the trend of the red line and blue line is the same, which means that the brightness estimation of using the button is comparably accurate. However, the accuracy of the result is worse than the first experiment. This can be caused by some predictable reasons. In the second experiment, the participant may be tired of concentrating on predicting the brightness. Since it's their second time to do the task, they may feel a little exhausted. For more time, the participant often underestimates the brightness, and the predicted result is often lower than the real result. Also, there is a possibility that the user forgot to change the brightness while they are doing the collecting skull task. According to Figure 1(c), the pattern of the user answer is corresponding to the true answer generally. The red line represents the true answer, and the trend of the red line and blue line is the same, which means that the brightness estimation of using the knob is comparably accurate. For the knob test, the accuracy of the prediction of the participant is better than the button. This means that when they use the Knob, it's easier for them to focus on the brightness of the sky. And also the Figure (d) shows the pattern of the user answer is corresponding to the true answer generally. The red line represents the true answer, and the trend of the red line and blue line is the same, which means that the brightness estimation of using the knob is comparably accurate. For the second experiment of the knob testing, the accuracy is not as good as the result of the first experiment at first, but it's better in the ending part. The figure 1 (e) shows the pattern of the user answer is corresponding to the true answer generally. The red line represents the true answer, and the trend of the red line and blue line is the same, which means that the brightness estimation of using the slider is comparably accurate. Compared to the knob test, the accuracy of brightness prediction is worse, but compared to the button, the result is better. At the first period of the experiment, it seems that the participants forget to change the prediction of brightness. But in the later part, it goes well. The last one of Figure 1 shows the pattern of the user answer is corresponding to the true answer generally. The red line represents the true answer, and the trend of the red line and blue line is the same, which means that the brightness estimation of using the slider is comparably accurate. For the second experiment of the slider test, the result is even worse. At first, it goes well and accurately predicts the brightness; however, in the middle part, the participant underestimates the brightness.

Also, from the three-position data, we can easily conclude that the Button test goes comparably far and it can show that when the participant uses Button, they can focus on the game but can't focus on the details. From the position data of the knob, the participant didn't go so far away from the original location. For the slider test, the distance movement is not too far. From the average score of the 3 tools, the score of the Button is higher than the score of the other tools, and the score of the Knob is higher than the score of the slider. When they use the Button, they focus more on the task of collecting the skull. When they use the slider, the score of the slider is the least, which means they spend more time on brightness prediction, instead of collecting the skull.

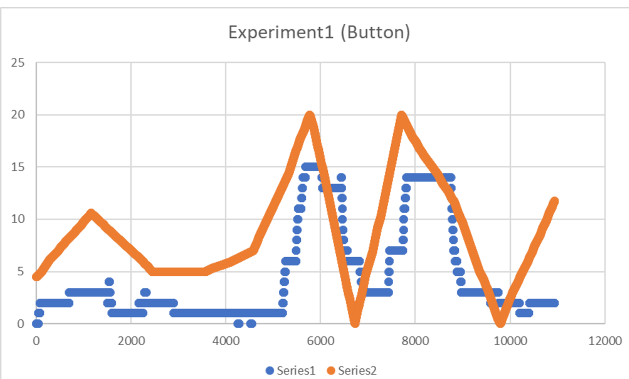

Figure 1 (a) Brightness Estimation (Button test in Experiment 1)

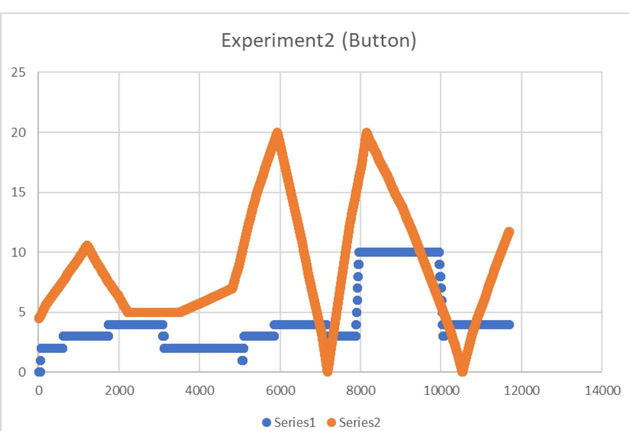

Figure 2 (b) Brightness Estimation(Button test in Experiment 2)

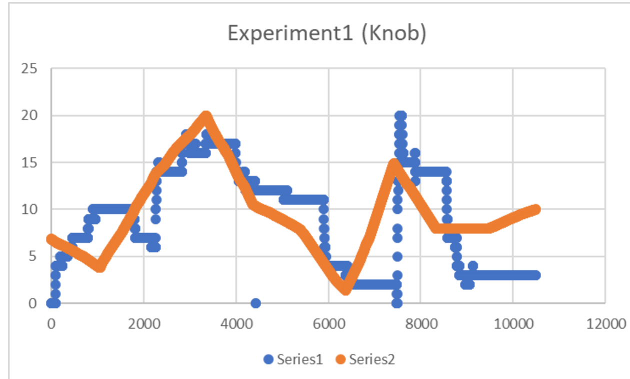

Figure 3 (c) Brightness Estimation (Knob test in Experiment 1)

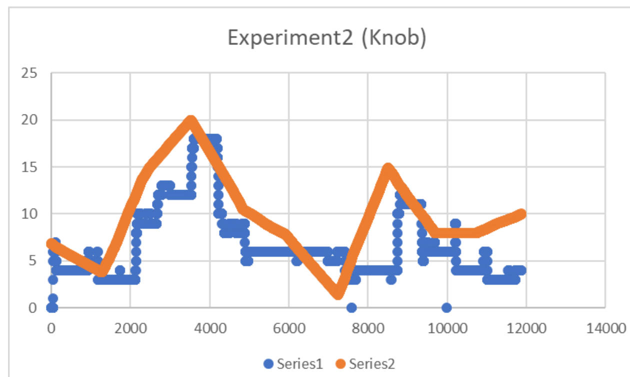

Figure 4 (d) Brightness Estimation (Knob test in Experiment 2)

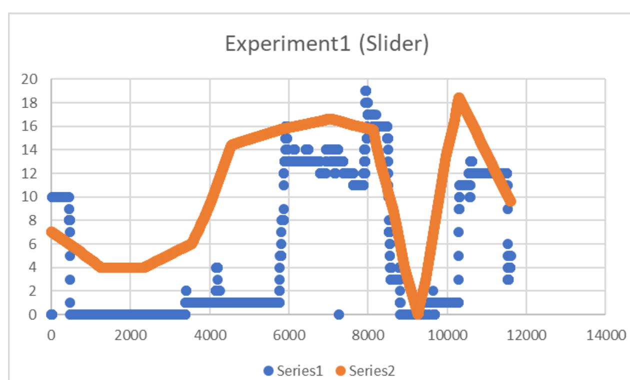

Figure 5 (e) Brightness Estimation (Slider test in Experiment 1)

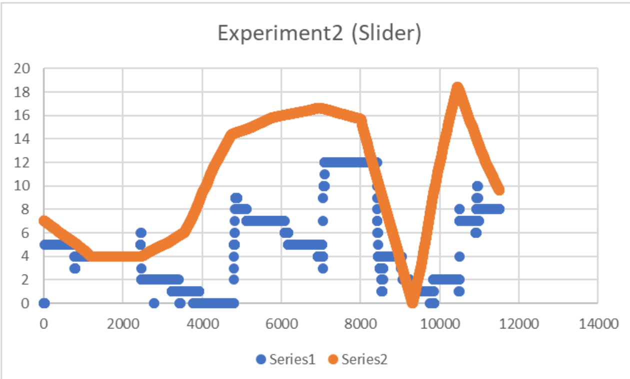

Figure 6 (f) Brightness Estimation (Slider test in Experiment 2)

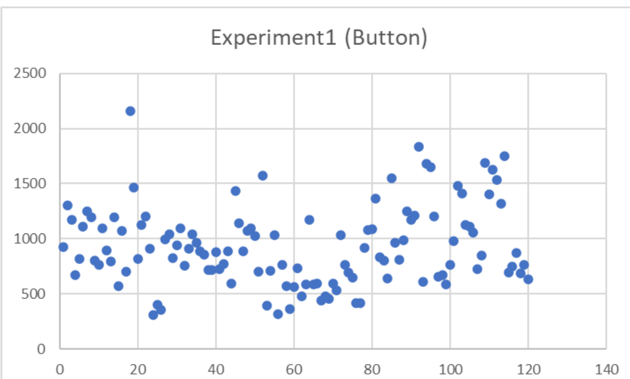

Figure 2 (a) Position data (Button test in Experiment 1)

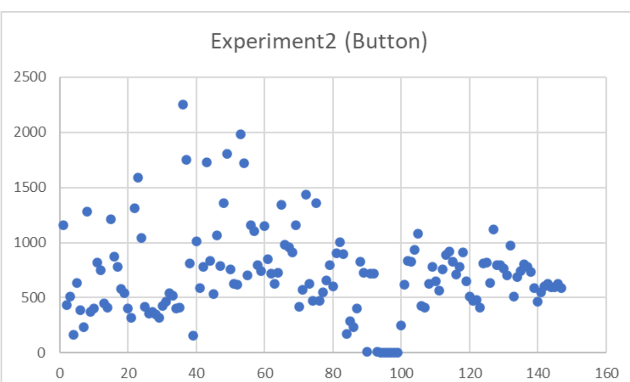

Figure 2 (b) Position data (Button test in Experiment 2)

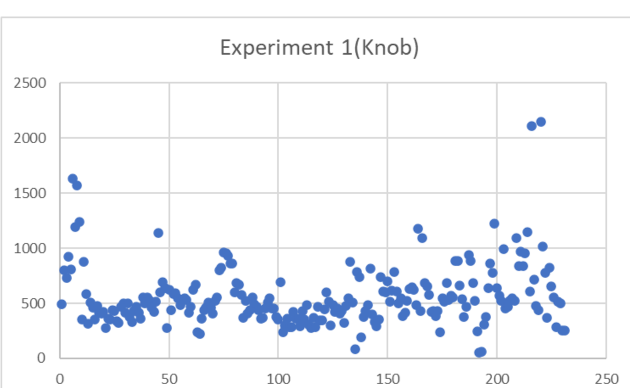

Figure 2 (c) Position data (Knob test in Experiment 1)

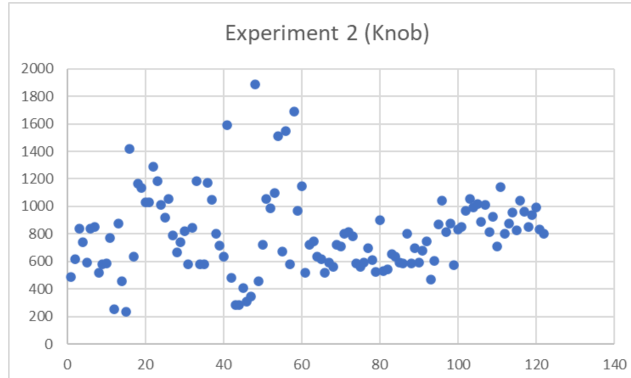

Figure 2 (d) Position data (Knob test in Experiment 2)

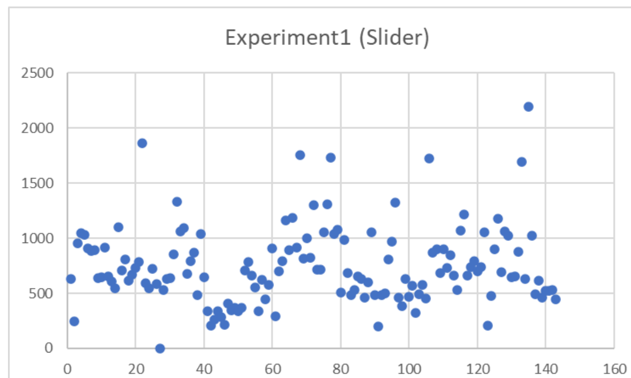

Figure 2 (e) Position data (Slider test in Experiment 1)

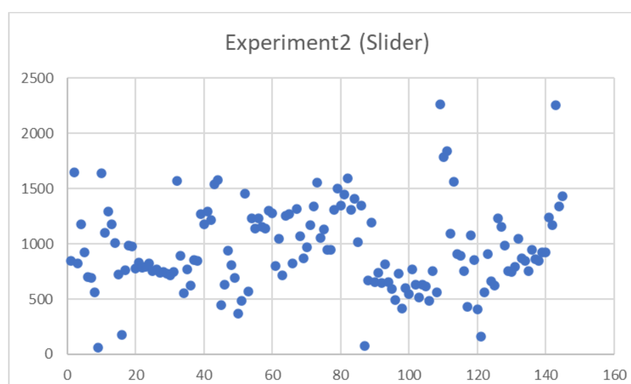

Figure 2 (f) Position data (Slider test in Experiment 2)

4. Conclusions

From the result of the experiment of the cybersickness test of Virtual Reality and also the process of comparing the three tools, we can conclude that each tool has its own benefit and deficiency and the most accurate one may be the knob, but the other two are also accurate in some degree. After we done this part of experiment, it is necessary for us to use the knowledge in the real cybersickness expression experiment. Further experiment needs to be implemented in which the winning device is used to gather cybersickness reports along with eye tracking and physiological data as the user is passively transported around the same virtual environment in a cybersickness-inducing manner (e.g. from the perspective of a kangaroo).

REFERENCES

McHugh, N., Hoermann, S., Jung S., Lindeman R. W. (2019). Investigating a Physical Dial as a Measurement Tool for Cybersickness in Virtual Reality. *VRST* '19 Parramatta, NSW, Australia.