An Empirical Study on Intuitive Gesture Manipulation in Virtual Reality

Conference Paper · May 2022

DOI: 10.1109/ICVR55215.2022.9848105

3 authors:


Jen-Tun Lee
Japan Advanced Institute of Science and Technology
11 PUBLICATIONS **6** CITATIONS
SEE PROFILE


Janaka Rajapakse
Tainan National University of the Arts
52 PUBLICATIONS **143** CITATIONS
SEE PROFILE


Kazunori Miyata
Japan Advanced Institute of Science and Technology
24 PUBLICATIONS **93** CITATIONS
SEE PROFILE

2022 8th International Conference on Virtual Reality

An Empirical Study on Intuitive Gesture Manipulation in Virtual Reality

Jen Tun Lee
School of Knowledge Science
Japan Advanced Institute of Science and Technology
Ishikawa, Japan
bmpsst511@gmail.com

R.P.C. Janaka Rajapakse
Graduate Institute of Animation and Film Arts
Tainan National University of Arts
Tainan, Taiwan
janakaraja@gmail.com

Kazunori Miyata
School of Knowledge Science
Japan Advanced Institute of Science and Technology
Ishikawa, Japan
miyata@jaist.ac.jp

***Abstract*—Gesture manipulation is a new input mode for the modern virtual reality (VR) user experience. Current gesture manipulation can be performed with three types of devices: Hand-held, Bare-hand and Hand-worn devices. Each device has its own advantages and disadvantages. However, many studies have not discussed the difference between these three types of gesture manipulation. To address this gap, we conducted research on how to find their learnability, effectiveness, and extensibility. We introduce a VR cup stacking game to mimic the state-of-presence experiment to compare the completion time of each device. Moreover, a series of user studies were conducted to collect feedback from participants. The results show that despite the Hand-held device being familiar to general users, but most of them were expected that the Hand-worn and Bare-Hand device can be well developed and wildly adopted into other applications to provide more intuitive and immersive experience.**

***Index Terms*—Free-hand interactions, Intuitive manipulation, Virtual Reality, Embodied operation**

I. Introduction

Virtual reality (VR) technology has been widely used in various fields, including education [1], [2], sports [3], [4] and entertainment [5], [6]. Interaction with virtual content plays a vital role in most VR experiences. An input device is a hardware tool in digital applications that allows users and programs to exchange data and interact with virtual content. For users new to VR, the most familiar characteristics of an input device are its shape and size due to the first look and feel of holding it. General interactive content is usually suitable for different input devices, and in a specific application, it is only suitable for input devices with particular usage patterns. Therefore, according to the main way of manipulation, choosing the appropriate input device is an important design decision.

We usually classify input devices according to their degrees of freedom(DOF). According to different applications, there are corresponding controllers from a single DOF to six DOF in VR. Standard VR input devices have six DOF (up&down, left&right, forward&backward, roll, pitch, and yaw), which allows the device to move and rotate in 3D space. In the tracking technology of VR-related equipment, the common ones are infrared tracking, inertial measurement unit (IMU), and camera tracking technology. The hybrid tracking system can combine the three advantages to provide a more intuitive operation experience for hand tracking. Currently, hand-held devices are used in primary use to perform these functions. A typical commercial hand-held device comes with buttons and can be converted into any indirect trigger by the user, and the button is also the most efficient and time-saving operation. On the other hand, like Microsoft Kinect, the developers of LeapMotion and VR Glove believe that too many buttons can confuse learning and operation errors, especially when the function design of the controls is inconsistent. Therefore, bare-hand, hand-worn, and hand-held devices have always been debated by different developers.

To this end, this research aims to comprehensively review the intuitive gesture manipulation provided by different VR input devices. We compared gesture manipulation in three types of device; using 1) handheld device (Vive Wands, ValveIndex), 2) bare hand device (LeapMotion) and 3) hand-worn device based on our previous work [7].

To better categorize each type of device, we addressed the following research questions.

• Learnability: We aim to evaluate the learning cost of each VR input device, and users can become familiar with pure hand interaction in a short time?

• Effectiveness: The aim of the VR input device is to provide users with a more intuitive and immersive experience, which is a standard for evaluating the effectiveness of each device.

• Extensibility: Which device has more potential and can be extended to other applications?

To verify each research question, we develop a VR cup stacking game for the participants and conduct a series of user studies, which will be explained in the following sections.

978-1-6654-7910-3/22/$31.00 ©2022 IEEE

Fig. 1. User Playing VR-cup stack game with different types of input devices.

II. Related Works

A. Intuitive Interaction

Intuition is an unconscious process and is difficult to verbalize to people. In the research by Brandenburg et al. [8], we realize that the concept of intuitive interaction is based on experience and can be observed by how people act with kitchenware or other objects in daily life. In most VR interaction cases, intuitive manipulation is faster than the conscious-thought mode [9]. Therefore, people intuitively use manipulations in situations they have often encountered before. Based on this understanding, we define an intuitive interaction as follows: Intuition is a cognitive process, a behavior that can be learned from experience, executed in the present moment quickly and unconsciously, and difficult to describe in words. People may not be able to explain how they make decisions during intuitive interactions [10].

Before, no designer knew how to apply everyday interactions to virtual reality to make it intuitive. We experimentally determined the relationship between intuitive interaction and familiarity and how different aspects of scene design affect intuitive interaction. The main findings of our study are as follows. Familiar users with similar features are faster at completing tasks and more able to perform interactive functions than users with less familiarity with related components.

B. Intuition Experiment

Intuition experiments have been proposed in the field of human-computer interaction as a way to test intuitive operation. Stefan Brandenburg and Katharina Sachse [8] designed an experimental environment in which they made a multitouch desktop (approximately 80 cm x 105 cm). The interface consists of a work area, a text description area, a start button, and three task objects on the right side of the work area. The task of the experiment was for the participants to perform three different actions (rotation, shearing, and scaling) on the object through gestures. They were evaluated two times: time to first click (TFC) and total task time (TTT). TFC represents the delay from pressing the start button to starting the gesture, while TTT represents the overall time from pressing the start button to completing the gesture task. The data showed that over time and with experience, subjects became faster in terms of the time required to perform gestures (TTT) and their initial reaction time (TFC).

Alethea L. Blackler, Doug Mahar, and Vesna Popovic [9]. In one of their experiments, they were asked to perform two actions, the first action is to use the camera's zoom function to take a picture in autofocus mode. The second action is to find and delete the just taken photo. Searches for the specified image stored in the camera and zooms in on it. The results showed that participants who had used similar functional products could complete the operation faster and more intuitively. The intuitive first-use results of the product are significant because, in the case of the first operation, the participants did not know where the specified function was, but they completed the task correctly in the first exploration. These are not physical affordances and, in most cases, are characteristics that are difficult to predict, so participants can only act on similar features they have manipulated in the past. Therefore, their results support the idea that having a similar action or the same text description in a product is more intuitive for people to use it for the first time.

Despina Michael-Grigoriou, Panayiotis Yiannakou, and Maria Christofi [11] conducted a study between groups with 22 participants equally divided between the two groups. Participants in the first "VR group" explored the skeletal system through VR gesture manipulation. In contrast, the second "SP group" (slide show presentation) explored the human anatomy through a slide show presentation. The above two methods can find the same information. Participants in both groups wrote a pre-test knowledge questionnaire (pre-KT) with a total score of 10 before experimenting. After experimenting, they then wrote the same knowledge questionnaire (post-KT). Furthermore, they gave participants in the VR group another questionnaire (5-point Likert scale) after the experiment, which was used to evaluate hand motion recognition techniques. From the results of their research in the first questionnaire, it can be seen from the data analysis of pre-KT and post-KT that VR technology has advantages over the slide presentation method. In the second questionnaire, participants in the VR group gave very high scores for the sense of immersion brought about by gesture operation, the cognitive sense of the system, and the ease of operation.

To study intuitive interactions in mixed reality game systems. Shital Desai, Vesna Popovic and Alethea L. Blackler [12] conducted an experiment in which 42 children aged 5 to 12 years were invited to play a mixed reality game system called Osmo from Tangible Play. They ran a related Friedman test to determine whether there was a statistically significant difference in the number of intuitive, non-intuitive, and partially intuitive interactions among children who experienced mixed-reality game systems. From their results, it can be concluded that intuitive interactions have higher average rankings than non-intuitive interactions and partial intuitive interactions.

C. *Spatial Presence Interaction*

This paper considers the presence of intuitive interactions in the context of VR, Augmented Virtual (AV), and MR. Martijn [13], mentioned that presence refers to the feeling of ”being in” a world, the process of when one environment (virtual environment) begins to overlap another (real environment). In other words: presence refers to the user’s distinction between RE and VE. However, many factors affect the sense of presence, such as story presence [14], cognitive presence [15], and affective presence [16], which will affect the VR experience of the user. Since it is beyond the scope of studies to consider all the presences, this paper focuses on the relevant definitions of spatial presence and interaction with three input devices.

III. Design Implementation

A. *Device Setup*

The user is immersed in the VR application by wearing a stereoscopic head mounted display (HTC Vive Pro) with position tracking and can be based on a gesture tracking device (Leap Motion) mounted on the HMD, or a Handheld Device (Valve Index & Vive wand) interact with the VR glove, as shown in Figure 2. Virtual Reality setup: To create a virtual reality environment, the user is immersed in the VR application by wearing a stereoscopic head-mounted display (HTC Vive Pro) with position tracking. In the PC hardware configuration, the processor is an Intel(R) Core i7-i7-8750H CPU, the highest is 4.10GHz, and the memory is 32GB. To provide a better visual experience and interactive effects for the experiments, we used a NVIDIA GeForce GTX 1070 8GB as a graphics card. Hand Tracking: We use three different types of devices for hand tracking. Hand-held Device (Valve Index & Vive wand): Vive wands have 24 sensors that can track movement in space through a lighthouse and a set of triggers to interact with virtual objects. The Valve Index controller features a joystick, a touchpad, two face buttons, a menu button, a trigger, and a set of 87 sensors that are capable of tracking the position of the controller in space, finger bending, movement, and pressure to simulate the finger movement of the user in VR as shown in Figure 2 (a & b). Bare-hand device: LeapMotion uses infrared scanners and monochromatic IR cameras to map and track the human hand. This information is used to create digital versions of indicators that can manipulate digital objects in real time. Therefore, we mounted LeapMotion in front of the headset so that the virtual hand of the user is always present in the scene, letting the user know that their hands are well tracked, as shown in Figure 2 (c). The hand-worn device is based on our previous research [7], can be divided into two parts, the glove and the data box; The glove part has five Flex sensors sewed to read the curl angle of five fingers. The data box includes a Wemos D1 mini, which is equipped with a WIFI module that allows us to implement wireless data transmission, an ads1115 capable of extending the analog pins to five pins, including the A0 pin on the MCU, and a custom PCB SMD resistor with five 10k Ohms to minimize box size; a Vive tracker was used to track the user’s hand position in space; the setup is shown in Figure 2 (d).

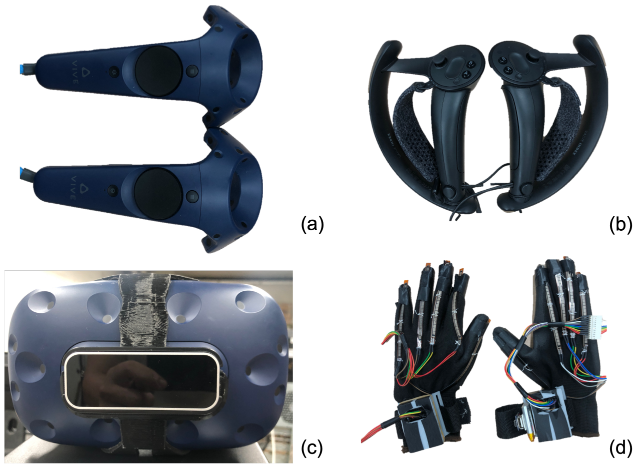

Fig. 2. Participants using Controller (a), ValveIndex (b), LeapMotion (c) and VR Glove (d) to perform the VR application.

B. *Grab Interaction Design*

Grab interactions have always been the typical way to interact with virtual objects on head-mounted VR devices such as Oculus Quest and HTC Vive [18]. For the handheld device, the SteamVR plugin [17] in Unity already provides basic interactive development functionality. Therefore, for the grab function of Vive wands and ValveIndex, we use the plugin to implement the grab interaction. For the bare hand device, UltraLeap provides Unity3D assets [18] for developers who intend to create content using LeapMotion. It contains essential grab interaction functions. For the hand-worn device, to make sure that the manipulation will not have too much different than above two devices. We refer to their grasping interaction design, which has the following utilities: object proximity detection, grab recognition, and object following. For example, for grasping objects, when the hand is close to the target to be grasped, the target’s color will change to bright red, which reminds the user that the hand is close to the effective distance. Then, the position of the target object will move with the palm when the grab gesture is detected.

IV. User Study

A. VR cup stacking game Design

This task requires the players to stack the cups regularly and restore them in the fastest time. For the scene part, we used Blender to design a table that complies with the Sport Stacking competition rule book [19], a table width: 72.5-77.5 cm (29-31 inches), length: 180-187.5 cm (72-75 inches), height: 72.5-77.5 cm (29-31 inches) to ensure that participants see the real size in the virtual world. We have changed some of the competition rules in VR. The original 3-3-3 project required 9 cups, each in a group of 3. Players must first stack each group of cups into a pyramid, and then reassemble them to return to their original shape. We let the user do not need to do the original final step. In place of that, there are three tasks, as shown in Figure 3, and one of them will be randomly assigned when participants press the start button. To complete the cup stacking task, they need to follow the order specified in the task image. To simulate a real case of intuitive manipulation, we set up two interactive modes: the 180° front test and the 360° surrounding test, which will be explained in this subsection.

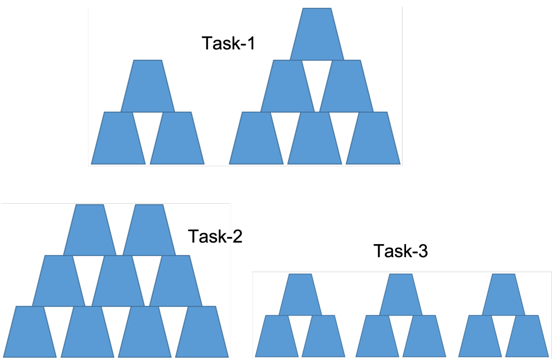

Fig. 3. Cup stacking tasks

B. State of Presence Experiment

Due to the design of the traditional console game, people get used to interact with the content present in front of them and play with a button-based controller. However, the state of presence is highly connected to the user experience in VR. Why VR can provide an immersive experience is because it can let people see the objects behind them [20]. Although VR already provides 360° content for users, no research discusses which type of controller is suitable for VR. We set up two state of spatial presence scenario to investigate the difference between the types of VR input devices.

1) First state of spatial presence scenario (FSP) (180°): In the 180° scenario, participants were placed in the center of a room simulating a cup stacking competition. The cups were already placed on the table and the participant's view was restricted to 180° of the room. The scene behind the participant was all black, as shown in Figure 4.

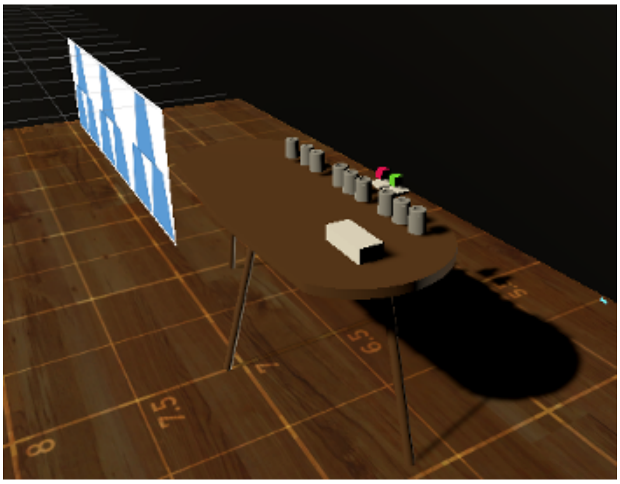

Fig. 4. Cups are aligned in front of user

2) Second state of spatial presence scenario (SSP) (360°): Unlike the first state of spatial presence scenario, the cups will appear randomly and surround the participant, forcing the user to look around and pick up the cups, as shown in Figure 5.

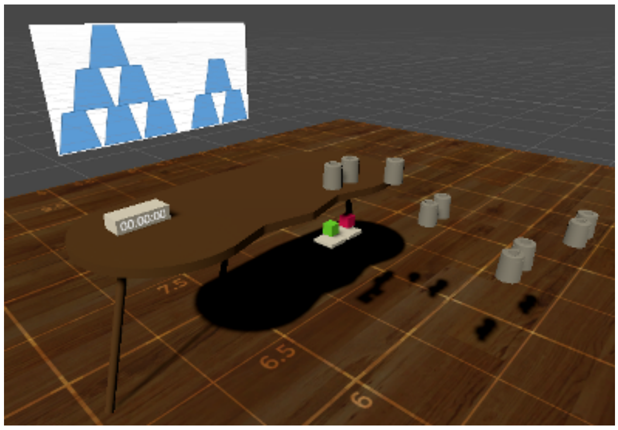

Fig. 5. Cups are surrounding user

C. Participants

We recruited 20 participants (8 males and 12 females) on campus through advertising to participate in our experimental study. The age of the participants ranged from 23 to 27 years (M = 24.5, SD = 1.51). We divided 20 people into two groups, each group consisting of 4 men and 6 women, to experience the first and second scenes, respectively. The participants were allowed to explore the application we developed before solving the task. For this group, it can be assumed that they have little experience with Hand-Worn devices and Bare-Hand devices (only two said that they had ever used LeapMotion to develop VR applications).

D. Procedure

Testing with participants who have not used VR controllers requires some special considerations. First, they do not know exactly how to use controllers, especially handheld controllers, so they need verbal instruction. Instructions must be presented in a standardized manner, as small changes in wording can significantly impact interactions. A male instructor read out the task to all subjects and first invited them to move their fingers or press the correspond button to activate the grab gesture in a virtual environment.

The first group of participants experienced the FSP. Before starting the operation, the participant must press the start button to start counting time, and the system will randomly assign the stacking task to the participant, as shown in Figure 3. When the cup stacking task is complete, press the red button to stop the timer. We let the participant rest for one minute and play the game again to collect the first and second results. The second group of participants experienced the SSP, and the procedure was the same as in the first scene. Participants who had experienced VR equipment were less attentive. To motivate them to complete the test, we told participants that this task is time-competitive, so do it as quickly as possible.

E. Results

To examine whether each type of input device is easy to learn, we collect the completion time, which is the time that the participants complete the tasks in the first T1 and the second T2, as shown in Figures 6 & 7. It shows us that when the participants use the same device for the second time, the time results are improving. The learning cost is obviously low on the handheld device. For example, users who use a handheld device (Means (M) of Vive wands T1= 35.9, T2= 22.9, ValveIndex T1= 38.1, T2= 28.9) are better than bare hands (M T1 = 78.4, T2 = 32.8) and a hand-worn device (M T1 = 41.1, T2 = 23.2), as shown in Table I & II. To verify the difference between each device, we conducted a MANOVA test as shown in Table III & IV. The results show that statistically significant differences in time performance are observed in four input devices tested in FSP, F(6,70) = 3.724, p <.003; Wilk's Λ = 0.575, partial η^2 = 0.242. In the SSP, F(6,70) = 5.918, p <.001; Wilk's Λ = 0.44, partial η^2 = 0.337. Scheffe's post hoc test was conducted for multiple comparisons. The results of the analysis is shown in Table III & IV. In the FSP test, LeapMotion took the longest time to complete the task (M T1 = 78.4s, T2 = 32.8s). Furthermore, if we look at LeapMotion in the T1 section, there is a significant difference between the handheld device (p_{Wands} = .002, $p_{ValveIndex}$ = .005) and the VR glove (p =.008). However, in the T2 section, LeapMotion has a similar time (M T2 = 32.8s) with other input devices and does not have a significant difference with Vive wands (p = .416), ValveIndex (p = .959) and VR Glove (p = .444). For the SSP test, LeapMotion still took the longest amount of time to complete the task in the T1 section (M T1 = 50.7s). However, VR Glove took the longest time in the T2 section (M T2 = 35.1s). Vive wands get the fastest time to complete the task (M T1 = 26.8 s, T2 = 22.9 s) and have the significant difference with VR Glove (p = 0.010), LeapMotion (p <.001) and ValveIndex (p = .121). However, in the T2 section, Vive wands and ValveIndex have similar time results (M $T2_{Wands}$ = 20.5 s, $T2_{ValveIndex}$ = 24 s), but faster than LeapMotion and VR Glove (M $T2_{LeapMotion}$ = 32.7s $T2_{Glove}$ = 35.1s).

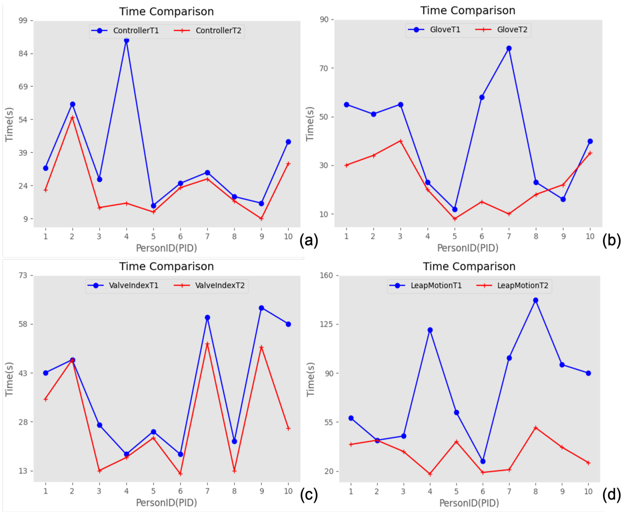

Fig. 6. T1 & T2 comparison in FSP

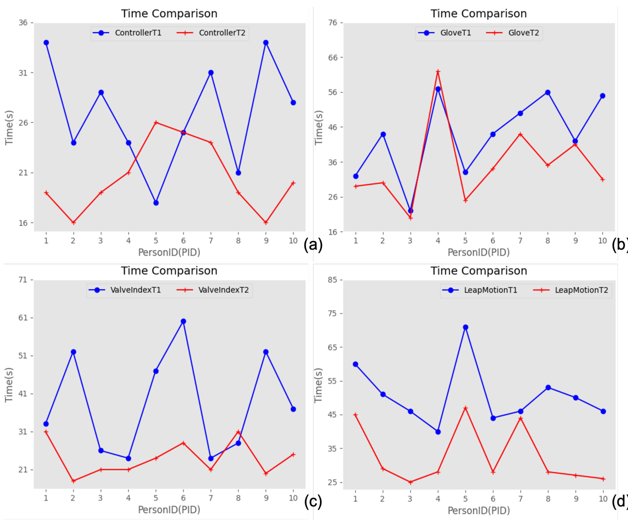

Fig. 7. T1 & T2 comparison in SSP

F. Questionnaire

The previous experimental results are similar to the results of Blackler et al. [21], that is, the use experience learned from other products can encourage users to quickly get used to the controller with similar functions. We found that appearance and perception style appeared to be the largest variables that affected the time spent on the task and the intuitive use of

TABLE I
DATA FOR THE COMPLETE TIME IN FSP

Group		Type of Input Device				Means
		Vive Wands	Valve Index	LeapMotion	VR Glove	
Participants	T1	32, 61, 27, 90, 15 25, 30, 19, 16, 44	43, 47, 27, 18, 25 18, 60, 22, 63, 58	58, 42, 45, 121, 62 27, 101, 142, 96, 90	55, 51, 55, 23, 12 58, 78, 23, 16, 40	
	T2	22, 55, 14, 16, 12 23, 27, 17, 9, 34	35, 47, 13, 17, 23 12, 52, 13, 51, 26	39, 42, 34, 18, 41 19, 21, 51, 37, 26	30, 34, 40, 20, 8 15, 10, 18, 22, 35	
Means	T1	35.9	38.1	78.4	41.1	48.38
	T2	22.9	28.9	32.8	23.2	26.95

TABLE II
DATA FOR THE COMPELETE TIME IN SURROUNDINGNESS (360°) TEST

Group		Type of Input Device				Means
		Vive Wands	Valve Index	LeapMotion	VR Glove	
Participants	T1	34, 24, 29, 24, 18 25, 31, 21, 34, 28	33, 52, 26, 24, 47 60, 24, 28, 52, 37	60, 51, 46, 40, 71 44, 46, 53, 50, 46	32, 40, 22, 57, 33 44, 50, 56, 42, 55	
	T2	19, 16, 19, 21, 26 25, 24, 19, 16, 20	31, 18, 21, 21, 24 28, 21, 31, 20, 25	45, 29, 25, 28, 47 28, 44, 28, 27, 26	29, 30, 20, 62, 25 34, 44, 35, 41, 31	
Means	T1	26.8	38.3	50.7	43.1	39.73
	T2	20.5	24.0	32.7	35.1	28.08

TABLE III
MULTIPLE COMPARISONS IN FIRST SCENE

DV	Type (I)	Type (J)	MD (I-J)	Std. Error	Sig.
T1	Vive Wands	Glove	-5.2	11.33	0.975
		LeapMotion	-47.3	11.33	0.002
		ValveIndex	-3.1	11.33	0.995
	Glove	Vive Wands	5.2	11.33	0.975
		LeapMotion	-42.1	11.33	0.008
		ValveIndex	2.1	11.33	0.998
	LeapMotion	Vive Wands	47.3	11.33	0.002
		Glove	42.1	11.33	0.008
		ValveIndex	44.2	11.33	0.005
	ValveIndex	Vive Wands	3.1	11.33	0.995
		Glove	-2.1	11.33	0.998
		LeapMotion	-44.2	11.33	0.005
T2	Vive Wands	Glove	-0.3	5.79	1.000
		LeapMotion	-9.9	5.79	0.416
		ValveIndex	-6.7	5.79	0.722
	Glove	Vive Wands	-0.3	5.79	1.000
		LeapMotion	-9.6	5.79	0.444
		ValveIndex	-6.4	5.79	0.749
	LeapMotion	Vive Wands	9.9	5.79	0.416
		Glove	9.6	5.79	0.444
		ValveIndex	3.2	5.79	0.959
	ValveIndex	Vive Wands	6.7	5.79	0.722
		Glove	6.4	5.79	0.749
		LeapMotion	-3.2	5.79	0.959

TABLE IV
MULTIPLE COMPARISONS IN SECOND SCENE

DV	Type (I)	Type (J)	MD (I-J)	Std. Error	Sig.
T1	Vive Wands	Glove	-16.7	4.61	0.010
		LeapMotion	-23.9	4.61	<0.001
		ValveIndex	-11.5	4.61	0.121
	Glove	Vive Wands	16.7	4.61	0.010
		LeapMotion	-7.2	4.61	0.495
		ValveIndex	5.2	4.61	0.737
	LeapMotion	Vive Wands	23.9	4.61	<0.001
		Glove	7.2	4.61	0.495
		ValveIndex	12.4	4.61	0.083
	ValveIndex	Vive Wands	11.5	4.61	0.121
		Glove	-5.2	4.61	0.737
		LeapMotion	-12.4	4.61	0.083
T2	Vive Wands	Glove	-14.6	3.54	0.003
		LeapMotion	-12.2	3.54	0.015
		ValveIndex	-3.5	3.54	0.806
	Glove	Vive Wands	14.6	3.54	0.003
		LeapMotion	2.4	3.54	0.927
		ValveIndex	11.1	3.54	0.032
	LeapMotion	Vive Wands	12.2	3.54	0.015
		Glove	-2.4	3.54	0.927
		ValveIndex	8.7	3.54	0.129
	ValveIndex	Vive Wands	3.5	3.54	0.806
		Glove	-11.1	3.54	0.032
		LeapMotion	-8.7	3.54	0.129

time during this period. However, when we started running the time quiz, we felt that these methods did not yet cover the relevant elements of the user's intuitive interactive experience provided by different devices.

We looked at different sources, such as the VR book [22], showing that questionnaires are often used to help evaluate interfaces that are already in use or have some operational component. On the Likert scale, which contains statements about a particular topic, participants indicated their level of agreement, ranging from strongly disagree to strongly agree. The options are balanced so that the number of positive and negative positions is equal.

Furthermore, the literature on intuitive decision making in psychology, the literature on human-computer interaction, and subjective reports of people's use experiences collected in interview studies [23]. Participants reported specific experiences in interacting with products they found intuitive and easy to use. Products cover a wide range (such as software, mobile phones, digital cameras, music players, home appliances, game consoles, answering machines, printers, navigation systems, copiers, etc.). They reflect on how they operate the product and the associated feelings and thoughts. After describing specific interactions, they also expressed a personal view of what a kind of intuitive interaction is and what a typical characteristic is.

In summary, to investigate intuitive and immersive feedback and the potential of each device, we decided to conduct a questionnaire to include various experiential features of intuitive interaction. With the above discussion, our questionnaire is based on the structure recommended in the VR book [22], and using a 5-point Likert scale, for VR applications and integrated hand motion recognition technology in immersive and intuitive interaction evaluation.

G. *Questionnaire Results*

Figure 8 & 9 shows the result from our questionnaire. Figure 8 shows the Likert scale scores that the participants rated their experience as more intuitive when they were using each device. Participants feel that their experience is more intuitive on the bare hand device. Figure 9 shows the scores of immersiveness and the scores of extensibility for each device. Participants felt that their experience is more immersive on the bare hand device, no matter on the FSP or SSP.

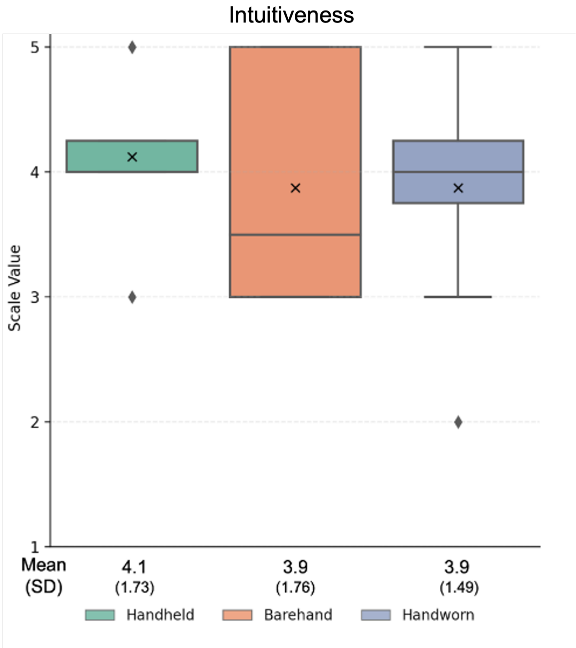

Fig. 8. Intuitiveness

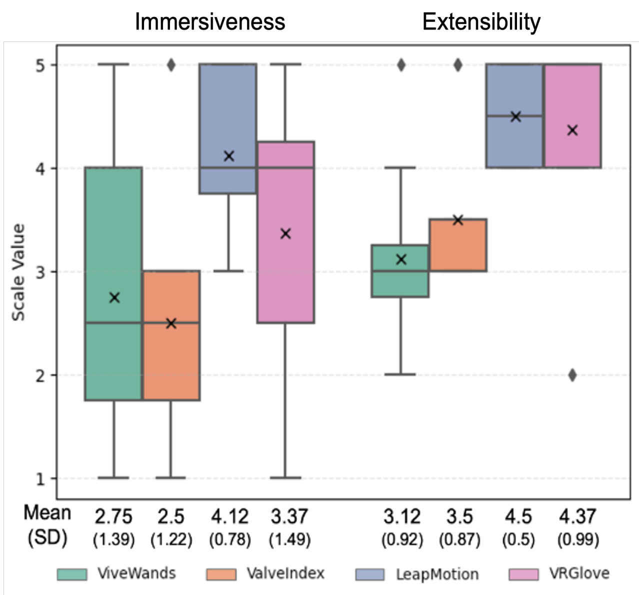

Fig. 9. Immersiveness and extensibility

V. Discussion

A. *Learnability*

In the FSP, the average times of T1 and T2 are quite similar for the handheld and hand-worn device, except the bare hand device, which takes longer than the others, as shown in Table 1 & II. That is because the participants are familiar with the button-based interaction from their daily experience. Bare hand and hand-worn device that is quite fresh for them to use physical hand gestures to interact with the digital content in the virtual world. P8 in group A reported that interaction through bare hand devices has the following problems. " When the cups are too scattered, I accustomed to grasping the target with both hands, but my sight does not follow the hand all the time, which will lead to hand tracking failed. When I stretch my hands, the camera has a limited range and cannot track my both hands well, significantly prolonging the time it takes for me to complete the task. " However, when the participants used LeapMotion for the second time, they learned how to avoid these limitations and spent less time than before, and the result is close to the handheld device. For the hand-worn device, P1 in group A said that " when he used it for the first time, the grasping works really smoothly in the virtual world as in the physical world. " We also can see the time result from Table I, which shows that the VR glove time result is close to the handheld device. There are also no statistically significant differences for each device, as shown in Table III. In the SSP, we expect that the time result will be longer than in the FSP. However, the result was not as long as in the FSP. P7 in group B said that " Although the cups are surrounding us, the distance is not too far from us. We still can grab them like what we do in the physical world. " In this case, we see that the performances of the bare hand and the hand-worn device are quite close to the ValveIndex. Apparently, the learning cost for the bare hand and hand-worn device is quite low for users who are first-time using it because pure hand interaction is what people do in their daily life.

B. *Effectiveness*

With any one of VR input devices, the participants give positive value to the intuitive experience. When experiencing our VR cup-stacking game, the participants prefer to use the handheld device, as can be confirmed in Figure 8. The reason is because this application is a competition, participants want to finish the task faster than the others.

When we ask the participants, "When I use the following devices, I can ignore the equipment (e.g. headset and gloves) and immerse in the virtual environment." P7 in group A said that " With the finger movements can be simulated, I will be more immersed and feel more realistic in the virtual environment. " In summary, the intuitive and immersive experience of each device depends on the task-oriented of VR application.

C. *Extensibility*

For the extensibility of the device, VR developers often strive to enable users to achieve an intuitive interactive experience with the virtual world through different input devices. In real life, one of the typical interaction methods, such as moving and clicking with a mouse or moving and pressing with a finger on a touch screen, has always been a traditional interaction method. This mapping method is easier to implement in flat interaction; however, this interaction mechanism is challenging to give users a real interaction experience in a three-dimensional space or without any instructions. In the three-dimensional virtual space, achieving natural interaction in reality is complex and dynamic [24]. In our questionnaire, when we asked participants to choose the device they thought could provide the most intuitive operating experience, there was not much difference between the four devices, but the participants said LeapMotion in addition to the recognition of grasping things, the response of finger simulation is excellent. The second-best device is the VR Glove. It can be seen that gestures play a critical role in improving immersion. Although in our experiments, users may not be able to quickly grab and release objects, raising their hands in the air can simulate a more realistic interaction, and most users believe that this type of device can be extended to other VR applications and help users become more engaged in virtual interactions.

VI. Conclusion

This paper investigates the effect of four devices on measuring the intuitive manipulation experience in the virtual environment through the VR cup-stacking game and questionnaires. Furthermore, the FSP and SSP shows us, that the cost of learning the devices of the bare hand and hand-worn types is very low for the users who are first-time using because pure hand interactions are what people do in their daily lives. This test contributes the insight of the potential that physical hand interaction can provide a similar manipulation experience as a handheld device does. In the questionnaire, the results have shown that the effectiveness of each device depends on the task-oriented of VR application. And the extensibility of each device, participants rate the bare hand and hand-worn device has the potential to provide a robust utility in other VR applications.

In our future work, we will optimize both bare hand and hand-worn devices. We expect to add more subjects and develop more interactive VR applications. The use of questionnaires to evaluate the level of intuitiveness and immersiveness may lead to inaccurate results due to the difficulty of quantifying such subjective experiences. We expect to use a brain wave device to collect the wave bandwidth during the experiment as objective data to analyze the level of intuitiveness and immersiveness.

Acknowledgment

Many thanks to the reviewers and editors for their valuable help and very constructive suggestions to improve this paper. We also would like to thank TNNUA students for their help in conducting the survey. Finally, we are thankful to all study participants whose reports and judgments on intuitive interaction provided an essential basis for the present research.

References

[1] M. Nooriafshar, R. Williams, and T. Maraseni, *The use of virtual reality in education*, 09 2004, pp. 6–8.

[2] G. Yildirim, M. Elban, and S. Yildirim, "Analysis of use of virtual reality technologies in history education: A case study," *Asian Journal of Education and Training*, vol. 4, pp. 62–69, 01 2018.

[3] C. McClure and D. Schofield, "Running virtual: The effect of virtual reality on exercise," *Journal of Human Sport and Exercise*, vol. 15, 12 2019.

[4] J. Qian, D. J. McDonough, and Z. Gao, "The effectiveness of virtual reality exercise on individual's physiological, psychological and rehabilitative outcomes: A systematic review," *International Journal of Environmental Research and Public Health*, vol. 17, no. 11, 2020. [Online]. Available: https://www.mdpi.com/1660-4601/17/11/4133

[5] C. Cruz-Neira, M. Fernández, and C. Portalés, "Virtual reality and games," *Multimodal Technologies and Interaction*, vol. 2, p. 8, 02 2018.

[6] D. Saffo, S. Di Bartolomeo, C. Yildirim, and C. Dunne, "Crowdsourcing virtual reality experiments using vrchat," 02 2020.

[7] L. J. Tun, R. J. Rajapakse, and K. Miyata, "Haptwarp: Implementation of bend, twist, and press manipulations for intuitive interaction in between the virtual and physical world," *Journal of Robotics, Networking and Artificial Life*, vol. 8, pp. 263–268, 2021. [Online]. Available: https://doi.org/10.2991/jrnal.k.211108.007

[8] S. Brandenburg and K. Sachse, "Intuition comes with experience," 10 2012.

[9] A. Blackler, V. Popovic, and D. Mahar, "Empirical investigations into intuitive interaction: A summary," *MMIInteraktiv*, vol. 13, 01 2007.
[10] ——, "Concepts of intuitive interaction," in *Workshop Proceedings der Tgung Mensch & Computer 2009*, H. Wandke, Ed. Germany: Logos Verlag Berlin, 2009, pp. 76–79. [Online]. Available: https://eprints.qut.edu.au/31031/
[11] D. Michael-Grigoriou, P. Yiannakou, and M. Christofi, "Intuitive interaction for exploring human anatomy in a vr setup," 10 2017, pp. 1–4.
[12] S. Desai, A. Blackler, and V. Popovic, "Intuitive interaction in a mixed reality system," 06 2016.
[13] M. J. Schuemie and C. A. van der Mast, "Presence: Interacting in vr?"
[14] S. B, I. Ladeira, C. Winterbottom, and E. Blake, "An investigation on the effects of mediation in a storytelling virtual environment," vol. 2897, 11 2002.
[15] D. Nunez and E. Blake, "Cognitive presence as a unified concept of virtual reality effectiveness," in *Proceedings of the 1st International Conference on Computer Graphics, Virtual Reality and Visualisation*, ser. AFRIGRAPH '01. New York, NY, USA: Association for Computing Machinery, 2001, p. 115–118. [Online]. Available: https://doi.org/10.1145/513867.513892
[16] C. Jicol, C. H. Wan, B. Doling, C. H. Illingworth, J. Yoon, C. Headey, C. Lutteroth, M. J. Proulx, K. Petrini, and E. O'Neill, "Effects of emotion and agency on presence in virtual reality," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411764.3445588
[17] V. Corporation, "Steamvr plugin."
[18] ULTRALEAP, "Ultraleap plugin for unity. https://developer.leapmotion.com/unity."
[19] W. S. S. Association, "Wssa sport stacking rule book," 2017, p. 10.
[20] M. Usoh, C. Alberto, and M. Slater, "Presence: Experiments in the psychology of virtual environments," 06 1999.
[21] A. Blackler, V. Popovic, and D. Mahar, "Investigating users' intuitive interaction with complex artefacts," *Applied Ergonomics*, vol. 41, pp. 72–92, 01 2010.
[22] J. Jerald, *The VR Book: Human-Centered Design for Virtual Reality*. Association for Computing Machinery and Morgan & Claypool, 2015.
[23] S. Diefenbach and D. Ullrich, "An experience perspective on intuitive interaction: Central components and the special effect of domain transfer distance," *Interacting with Computers*, vol. 27, 04 2015.
[24] H. J. Kang, J.-h. Shin, and K. Ponto, "A comparative analysis of 3d user interaction: How to move virtual objects in mixed reality," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2020, pp. 275–284.