Alternative Design For An Interactive Exhibit Learning In Museums: How Does User Experience Differ Across Different Technologies-VR, Tangible, And Gesture.*

*Note: Sub-titles are not captured in Xplore and should not be used

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

***Abstract*—This paper investigates three types of user interfaces: VR, Gesture-based interface, and Tangible-based interface. We examine how user experience differs across different technology and what are the factors that make the experience difference. To find the answer we conduct an empirical study, in which we create three different interactive exhibits that apply these technologies to deliver the same scientific content about biotoxin in nature. The study uses a mixed-method, qualitative and quantitative, and measures two factors. First, user experience is measured by six dimensions of user experience: attractiveness, perspicuity, efficiency, dependability, simulation, and novelty. Second, attention holding power is measured by playing time. The study uses the semi-structured interview to emphasize the issue and learning media of each interface. Thirty-one subjects joined the study. The statistical results shows that there are significantly different user experiences when using a different type of interface. There are difference across five user experience dimensions, only novelty is relatively unchanged. Difference are primarily between VR and Gesture, and Tangible and Gesture. There is no significant difference in holding power between the three types of interface. The statistical analysis of result and interview feedback from participants suggest six aspects to focus on when choosing an alternative interface to create a new interactive exhibit: the novelty, user-friendly, precision of the input device, task and device design, multimodal of feedback, and quality of text in VR.**

***Index Terms*—Virtual reality, Interactive systems, User experience, Human computer interaction, Museum learning**

I. Introduction

There are many choices of technology for creating new interactive exhibits, which can make choosing the most effective technology challenging for exhibit designers. Different technologies, like virtual reality (VR), interactive displays, and tangible interfaces can deliver the same exhibit content, but the choice of technology may have impacts on exhibit attractiveness, user engagement, and learning outcomes. This paper presents a study comparing how technology choice influences these factors.

Previous work has explored how exhibit content can be delivered using different interaction technologies. VR is an immersive technology which provides visual and audio content to create a presence in a virtual environment. For example, dioramas and animal habitats are classic media that museums could deliver using VR. Although visitors may lose social interaction while in VR, the benefit of using VR is that it gives museums more flexibility to show a wide variety of content. To demonstrate that VR has benefits over other technologies for museum exhibits, this paper aims to investigate factors that influence user experience when using a different type of interface and factors that we should consider when choosing technology to develop an interactive exhibit for learning in a STEM museum.

We conduct an empirical study to compare user experience between different types of user interface. The study chooses three types of interface, VR, Tangible-based interface, and Gesture-based interface, to create three different interactive exhibits. Exhibits differ in terms of manipulating (input device) and display device, but use the same scientific content, narrative style, action and activity. All of the exhibit designs are based on experiential learning and make the exhibit feel immersive by simulating the natural environment with audio and visual. The study uses mixed method quantitative and qualitative research, gathering data from user feedback in a questionnaire and semi-structured interview. Analysis of the results suggest six aspects to consider when choosing technology to create a new interactive interface.

II. Background

A. *Benefits of Immersion for Learning in Museums*

The immersion experience seems to refer to a feeling of being in a time and place such as an historical period, animal habitat, a geological formation, or space flight. The focus of the immersion experience is on the experience rather than on formal learning [1].

A relative factor which impacts on immersion experience suggested by Bitgood [1] includes realism of the illusion, dimensionality, multi-sensory stimulation, meaningfulness, mental imagery, and lack of interfering factors. The immersion experience in the museum is more fun and exciting than a formal learning experience. Learning through the exhibit designed with a immersive experience is more memorable than a reading only approach. Gilbert [2] interviewed museum professionals about using the immersive approach for exhibition, finding three main reasons. First the immersive exhibition can attract people as their a leisure-time activity that can compete with other business. Second, they have a good holding visitors' attention and offer a memorable experience. Third, they deliver the content of the exhibit effectively.

B. Interaction Styles for Museum Exhibits

Interactive exhibits in museums can be classified into two types, hand-on interactive and interactive multi-media (IMM). Hand-on refers to exhibits that visitors can use their body, especially their hands, to interact with or manipulate the physical exhibit. IMM refers to multimedia that use a computer to control with interaction by touching the screen, push-button, mouse, and keyboard [3].

The character of the interactive exhibit in the museum is edutainment style, which entertains and educates visitors at the same time. There are typically many interactive exhibits in the same museum, with visitors free to choose to play with the interactive exhibits, often with limited time to visit the museum, so that interactive exhibits in museums need to provide a stimulus to continue in the first ten seconds [4].

Interactive exhibits in museums are developed for the museum to provide an edutainment environment. Packer [5] said that learning in museums is a learning experience for fun, surrounded by a combination of discovery, exploration, mental, stimulation, and excitement. Learning for fun has four conditions: a sense of discovery or fascination, appeal to multiple senses, the appearance of effortlessness, and the availability of choice. Interactive exhibits in museums consist of a short time playing in an interactive style in combination with delivering knowledge content.

C. Comparing Modalities for Museum Exhibits

The Hand-off exhibit does not allow visitors to touch the object, in contrast to the hand-on exhibit allows visitors to have an experience and obtain an understanding from an exhibit by touching objects [3]. To promote experiential learning or learning by doing, hand-on is an important tool for museums. Visitors quickly understand to interact with a physical object, as the perceive of a concrete physical object is a cue to interact with it.

The tangible interface is simples than other types of interface. The visitor is familiar with interacts naturally with the tangible interface [6]. A tangible-based interface allows social interaction or collaborative learning, as it can be multiple players [7], [6]. The limitation of the tangible object as a learning media is that it cannot display much information, only its physical self deliver scientific content. An advantage of technology nowadays is that a tangible interface case be used as input of the system, to activate content display more information in digital format. Tangible-based interface benefit from a combination of the concreteness of the physical object and abstract digital representations to help the learner more easily to understand new or difficult content in a more relaxed learning environment [8], [6].

On the other hand, the gesture-based interface is a touchless interface, mainly using the human body as an input device [9]. Users interact with the system by moving parts of the body, especially a hand or the head. The gesture-based interface is a natural user interface that allows people to interact with the system the same as they interact with the physical real world, such as the use of their voice, their hand, and their body [10]. Gesture-based interface is grounded on theories of embodied cognition. The human mind is closely connected to the sensorimotor experience, body activity influence the human mind, so the gesture-based interface has the advantage of enhancing body-related experience [9]. The interface allows player understanding, by perceiving the world by their body movement.

Virtual Reality(VR) usually uses Head-Mounted Display(HMD) to display the 3D environment and the viewer will wear a VR headset to have an immersive experience. Many museums adopt VR in their museum for immersive learning, which has benefits of immersing the learner into learning content in a 3D virtual environment [11]. Learning with VR increases enjoyment and concentration compared with non-immersive video display [12]. In general, the controller is the input device for VR. A controller is a tangible object that makes a virtual game (action in the virtual environment) tangible [13].

For interaction with an object in the virtual environment, naturalness is a key factor that to improve intuitive interaction. Naturalness occurs when physical control and virtual output has directional concordance. Increasing natural mapping control enhance the familiarity of the player to use the system [13], [14]. The study in [14] showed that incomplete tangible mapping (e.g.. VR controller) and realistic tangible natural mapping has higher familiarity than kinesics natural mapping (e.g. gesture).

It seems that VR has more positive aspect for learning than tangible or gesture. The design feature of VR technology provides a controller that might simulate a sense of touch and make the system intuitive with the natural movement of the VR controller in the virtual world. The disadvantages of VR include the possibility motion sickness and hygiene, especially for museum learning also VR does not support multiple users allows shared experience with other visitors. This paper will compare user experience between these interfaces and holding power that indicate encouragement of an exhibit. The study has two hypotheses. H1: There is no difference between the user experience quality when using a tangible-based interface, a gestures-based interface, and a VR interface. H2: VR has a higher holding power than the gesture-based interface and the tangible-based interface.

III. METHODS

A. Study design

The study chooses three alternative technologies. The focus of the study is VR. We choose two alternative interfaces, with differing features, to provide comparison and contrast . All the interactive exhibit deliver the same scientific content and have the same drag and drop activities. The chosen interfaces are VR, Tangible and Gesture. The study use within-group study by recruiting 30 people to join the study. We expect to use the qualitative data from 10 people who have played with each interactive exhibit as the first order in a series assigned by the researcher.

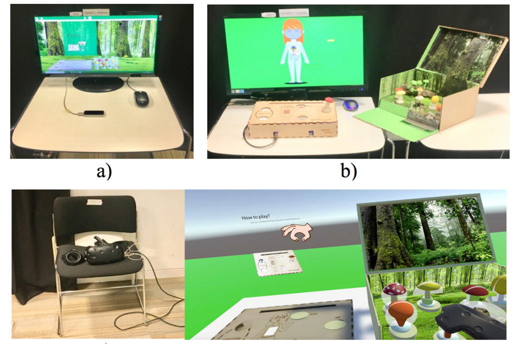
The study uses qualitative and quantitative data. The quantitative data measures two factors, sufficient user experience and holding power. The sufficient user experience is measured by using UEQ (User Experience Questionnaire) with seven scales on 26 questions [15]. The results are summarized into six criteria: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. The holding power is measured by the time that people play with each interactive exhibit. Qualitative research will use a video recorded, semi-structured interview explore issues of each interactive interface.

The experiment takes approximately 30-40 minutes. Three of the interactive exhibit are set up inside the room before the participant arrives. The participant signs a consent form and answer the questions in the questionnaire, including questions about their general background and experience with each type of interface. Using the Latin Square, the researcher assigns a random order to the three interactive exhibit for the participant to play with. The participant plays with the interactive exhibits one by one. After the participant finishes playing with each exhibit, they are asked to complete the UEQ and interviewed them about their experiences interacting with each interactive exhibit. At the end of the experiment, after they have experienced all three types of interactive exhibits, they will be asked about their overall experience.

B. Content and Narrative Story

The scientific content of the exhibit tells a story about biotoxin plant and animal in nature. The story includes three poisonous mushrooms, two edible mushrooms, one normal frog and one poisonous frog. The presentation is open exploration without a mission to complete. The instruction invites the player to explore by picking up an object and putting it on an area that represents eat, touch and smell, and get a feedback from the system. The touch, smell and eat areas are analogous to what happens when humans eat, touch or smell the object. The model of objects shows their anatomy

The feedback of the system is multimodule feedback including a warning colour, sound beep and action of animation. There are 24 different animations give information by showing



c)

Fig. 1. Examples of interactive exhibit used in the study. a) Exhibit 1: Gesture-based interface b) Exhibit 2: Tangible-based interface c) Exhibit 3: VR

the effect on a human organ, and telling the scientific name and common name of an object. There are three different warning colours; red means dangerous(poisonous), yellow means should consider, and green means safe. There are two different sound that indicate safe and dangerous things. Fig. 2 shows an example of the feedback from the interactive exhibit.

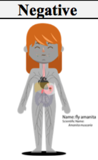
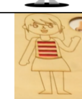
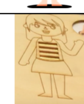
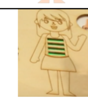
Modality	**Meaning**		
	Negative	**In between**	**Positive**
Visual Animation			
Visual Colour lights			
Audio	**Three tone beep**	**none**	**Single beep**

Fig. 2. Examples of the feedback on the interactive exhibit

C. System design

Three types of interactive exhibit interface have a specific technology that the researcher chose to create the systems. We use Unity3D as the main software to develop each interface. A detailed description of the technology used for the development of each interface as follows:

1) Exhibit 1: Gesture-based interface: UI01-Gesture,this interface uses Leap Motion as the input of the system and shows feedback on a PC screen. Users can move their hand over the sensor and can pick up an object on the screen by moving their hand over the object and closing their hand to pick up the object. The user can drop the object on the target area by opening their hand. All activities and the feedback of the system are shown on one screen.

2) Exhibit 2: Tangible-based interface: UI02-Tangible, the system includes two parts, a digital display and a physical display. The digital part shows the animation and information feedback on a PC screen. The physical part is an input of the system along with feedback shown on the physical box.

It includes an interactive play box, the six mushroom model, and two frog model. The interactive box developed by using a microcontroller, including two Arduino Mega 2560 boards and five Arduino Nano connect to RC522 RFID 13.56MHz. module. The mushroom and frog model have a NFC tag attached underneath. A LED is used to light up to show the status when the user moves an object and drops it inside the target area on the interactive play box.

3) Exhibit 3: VR-based interface: UI03-VR, the system uses an HTC VIVE VR Headset. Users can play with the system by using an HTC VIVE controller. Users can press the Trigger button when they want to pick up an object in the scene, and the user can drop an object by releasing the Trigger button. The feedback of the system is divided into two areas: the virtual interactive play box on the virtual table in the virtual environment, and an animation appear in front of the players.

IV. Results

Thirty-one people participated in this study, including seventeen male, thirteen female, and one other gender. Twenty-two people have visited the Science and Technology museum before. Eighteen have never experienced using a gesture-based interface. Twenty-two do not have experience playing with a tangible interface. Nineteen have experience using VR. Twenty-one reported being unfamiliar with background knowledge of Biotoxin in nature story.

A. Exhibit Technology and User Experience

Each item on the UEQ is transfered to a scale from -3 to +3. The +3 means very positive and the -3 means very negative. Data from the UEQ are divided mean values into three categories: negative evaluation $(values < -0.8)$, neutral evaluation (values between -0.8 to 0.8), and positive evaluation $(value > 0.8)$.

1) Exhibit 1: Gesture: Results for user experience of the Gesture-based interface show eleven positive evaluations, $thevalues > 0.8$ include: conservative/innovative (M = 1.68, SD = 1.14), dull/creative (M = 1.68, SD =1.17), conventional/inventive (M = 1.58 , SD = 0.96), not understandable/understandable (M = 1.45, SD =1.92), unattractive/attractive (M = 1.32, SD = 1.08), not interesting/interesting(M = 1.29, SD= 1.32), not secure/secure (M = 1.27, SD = 1.68, N = 30), unfriendly/friendly (M = 1.06, SD = 1.59), boring/exciting (M = 0.97, SD = 1.49), cluttered/organized (M = 0.87, SD = 1.67), inferior/valuable (M = 0.81, SD = 1.22). The others fifteen items received a neutral evaluation. Results are shown in Fig. 3.

2) Exhibit 2: Tangible: Results for user experience of the tangible-based interface show that almost all items received a positive evaluation, $thevalues > 0.8$. There are only two items with a neutral evaluation, the value between -0.8 to 0.8; usual/leading edge (M = 0.45, SD = 1.52), and slow/fast (M = 0.23, SD = 1.54). The top five highest mean values are for unfriendly/friendly (M = 2.39, SD = 0.84), complicated/easy (M = 2.23, SD = 0.88), difficult to learn/easy to learn (M = 2.13 , SD = 0.96), not understandable/understandable (M = 1.97, SD = 1.20), and inferior/valuable (M = 1.87, SD = 1.09). Results are shown in Fig.4.

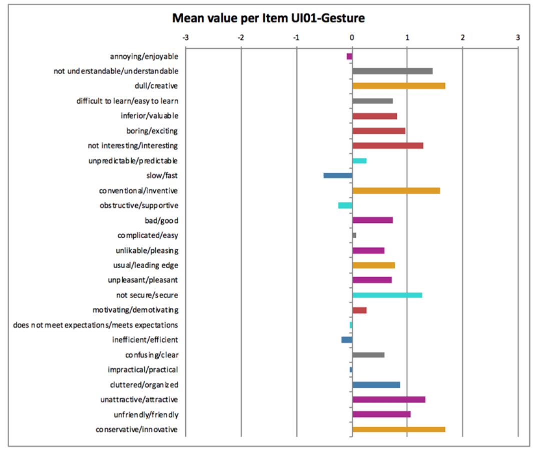


Fig. 3. The mean value per item of UI01-Gesture-based interface.

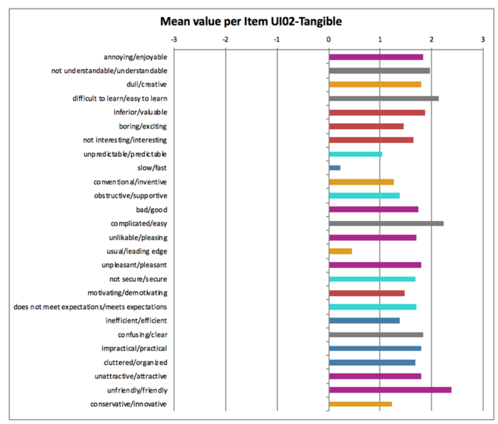


Fig. 4. The mean value per item of UI02-Tangible-based interface.

3) Exhibit 3: VR: Results for user experience of the VR interface show that only one item received a neutral evaluation; slow/fast (M = 0.74, SD = 1.59), and all other items received positive evaluation. The top five highest mean values are unfriendly/friendly (M = 2.35, SD = 0.71), difficult to learn/easy to learn (M = 2.19, SD = 0.91), cluttered/organized (M = 2.13, SD = 1.12), annoying/enjoyable (M = 2.06 , SD = 1.00), and inferior/valuable (M = 2.06 , SD = 1.03). Result are shown in Fig.5.

B. Factors of User Experience

The statistical results come from using the Friedman non-parametric test followed by a post hoc analysis with Wilcoxon signed-rank tests to analyse the different between pairs of technology. Post hoc Wilcoxon signed-rank tests were conducted with a Bonferroni correction applied, resulting in a significance level set by $\rho < 0.017$ (α = 0.05).

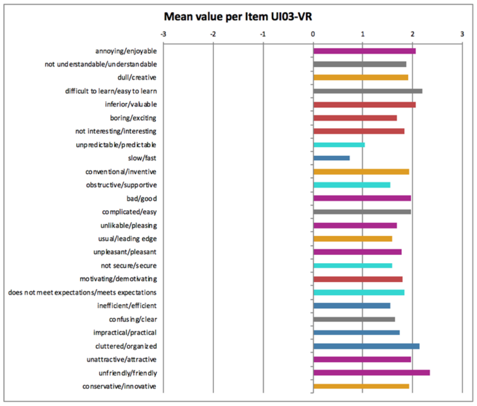


Fig. 5. The mean value per item of UI03-VR.

1) Attractiveness: There was a statistically significant difference in user experience on attractiveness depending on the type of interactive exhibit interface, $\chi^2 = 26.235, p = 0.000$. Median(IQR) attractiveness levels for the UI01-Gesture, UI02-Tangible and UI03-VR interactive exhibit running test were 0.40 (-0.40 to 1.60), 2.0 (1.40 to 2.60), and 1.75 (1.50 to 2.25), respectively. There were no significant differences between UI03-VR and UI02-Tangible based interface running trials (Z = -1.106, p = 0.269). However, there were significant differences between UI03-VR and UI01-Gesture-based interface running trials (Z = -4.018, p = 0.000) and between UI02-Tangible and UI01-Gesture-based interface running trails (Z = -3.861, p = 0.000).

2) Perspicuity: There was a statistically significant difference in user experience on perspicuity depending on the type of interactive exhibit interface, $\chi^2 = 21.876, p = 0.000$. Median(IQR) perspicuity levels for the UI01-Gesture, UI02-Tangible and UI03-VR interactive exhibit running test were 0.50 (-0.50 to 2.00), 2.25 (1.25 to 3.00), and 2.00 (1.25 to 2.75), respectively. There were significant differences between UI03-VR and UI01-Gesture based interface running trials (Z = -3.008, p = 0.003) and between UI02-Tangible and UI01-Gesture based interface running trails (Z = -3.804, p = 0.000). However, there was no significance between UI03-VR and UI02-Tangible based interface running trials (Z=-0.385, p = 0.700).

3) Efficiency: There was a statistically significant difference in user experience on efficiency depending on the type of interactive exhibit interface, $\chi^2 = 26.248, p = 0.000$. Median(IQR) efficiency levels for the UI01-Gesture, UI02-Tangible and UI03-VR interactive exhibit running test were 0.00 (-1.00 to 0.75), 1.5 (0.75 to 1.75), and 1.50 (1.00 to 2.50), respectively. There were significant differences between UI03-VR and UI01-Gesture based interface running trials (Z = -4.159, p = 0.000) and between UI02-Tangible and UI01-Gesture based interface running trails (Z = -3.996, p = 0.000). However, there was no significance between UI03-VR and UI02-Tangible based interface running trials (Z=-1.158, p = 0.247).

4) Dependability: There was a statistically significant difference in user experience on dependability depending on the type of interactive exhibit interface, $\chi^2 = 31.113, p = 0.000$. Median(IQR) dependability levels for the UI01-Gesture, UI02-Tangible and UI03-VR interactive exhibit running test were 0.50 (-1.00 to 1.500), 1.50 (1.00 to 2.25), and 2.00 (1.50 to 2.75), respectively. There were significant differences between UI03-VR and UI01-Gesture based interface running trials (Z = -4.080, p = 0.000) and between UI02-Tangible and UI01-Gesture based interface running trails (Z = -3.570, p = 0.000). However, there was no significance between UI03-VR and UI02-Tangible based interface running trials (Z=-1.701, p = 0.089).

5) Stimulation: There was a statistically significant difference in user experience on stimulation depending on the type of interactive exhibit interface, $\chi^2 = 30.154, p = 0.000$. Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set by p ¡ 0.017. Median(IQR) stimulation levels for the UI01-Gesture, UI02-Tangible and UI03-VR interactive exhibit running test were 0.75 (0.00 to 1.500), 1.50 (1.25 to 2.50), and 2.00 (1.25 to 2.75), respectively. There were significant differences between UI03-VR and UI01-Gesture based interface running trials (Z = -4.171, p = 0.000) and between UI02-Tangible and UI01-Gesture based interface running trails (Z = -3.437, p = 0.001). However, there was no significance between UI03-VR and UI02-Tangible based interface running trials (Z=-1.200, p = 0.230).

6) Novelty: There was not a statistically significant difference in user experience on Novelty, and it does not depend on the type of interactive exhibit interface, $\chi^2 = 0.475, p = 0.789$. Median (IQR) novelty levels for the UI01-Gesture, UI02-Tangible and UI03-VR interactive exhibit running test were 1.40 (0.800 to 2.20), 1.40 (0.60 to 1.80), and 1.80 (1.2 to 2.20), respectively.

In summary, the statistical information running test of each dimension shows us there is a difference in user experience when using a different type of interactive exhibit interface. There is an effect on attractiveness, perspicuity, efficiency, dependability and stimulation dimensions between using UI01-Gesture and UI02-Tangible, and in between using UI01-Gesture and UI03-VR. However, these five dimensions show no statistically significant difference between using UI02-Tangible and UI03-VR. On the other hand, the Novelty dimension shows no effect on user experience between the three types of interactive interface. In-depth detail of the different experiences between each type of interactive exhibit interface discussed in the analysis and discuss section.

C. Exhibit Technology and Holding Power

Holding power is a criteria for measure success of an exhibit [16]. Holding power or viewing time refer to the amount of time that visitor stop to play with exhibit, the time that visitors receive knowledge message from the exhibit, time visitor play with the interactive exhibit. VR is a immersive technology

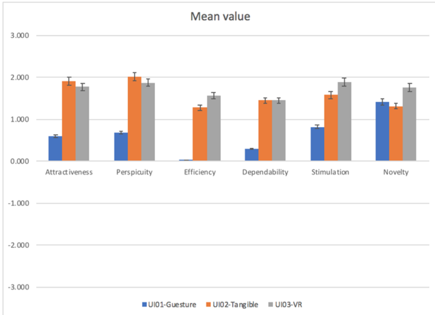


Fig. 6. show overall mean value per user experience dimension on UEQ for each interactive exhibit..

might have holding power than other technology. The statistically from this study show that there was not a statistically significant difference in playing time which indicates holding the power of the interactive exhibit. This does not depend the type of interactive exhibit interface, $\chi^2 = 3.528, p = 0.171$. Median (IQR) playing time in second for the UI01-Gesture, UI02-Tangible and UI03-VR interactive exhibit running test were 206.00 (131.00 to 291.00), 177.00 (139.00 to 214.00), and 180.00 (143.00 to 240.00), respectively.

V. Analysis and Discussion

The study design within-group study that a participant will repeat play with the same content on three different exhibit interfaces, so to access learning will use user experience feedback from ten people who assigned to play with each exhibit at the first order. The statistics show that it is different in user experience between VR and Gesture, and Tangible and Gesture, but no difference between VR and tangible on attractiveness, perspicuity, efficiency, dependability, stimulation. The novelty does not affect user experience in three different interfaces. All type of interface in this participant reported that they feel immersive like in the forest with sound background and forest scene. In-depth detail of user experience feedback will be discussed in this section.

A. Key Features

Across all three exhibits, participants gave positive feedback on the visuals and animations. It was fun and friendly. Almost all participants stated that the sound background is immersive across three exhibits. It makes them feel like they are on a walk in the forest.

The novelty was not statistically significant in the UEQ comparison, but participants talked about how new input techniques could be interesting. However, in VR four of the participants said they liked using the controller to pick up the object, and in gesture-based three of the participants mention they liked to see their hand on the screen and play with their hand. For museum exhibits, understanding the trade-offs between unfamiliar and familiar input techniques is important for user experience.

B. Interaction Challenges

Almost all participants stated that tangible-based interface and VR were easy to use. Picking up and dropping an object enjoyable using the system. On other hand, most of them said the gesture-based interface was difficult to move an object. We found that the sensor often did not detect the player's hand, makng it difficult for them to reach an object in a virtual scene and causing annoyance to use the system. One participant said, "…my hand is hurt.." The consequence of difficulty is that the player stops interacting with the exhibit.

Corresponding statistical results show that perspicuity, efficiency, and dependability varied significantly between tangible-based and gesture-based , and between VR and gesture, with no difference between tangible interface and VR. To explain this we refer to Task-Technology Fit (TTF) [17]. The basic TTF model involves task characteristics and technology characteristics. Good matching between the characteristic of a task and the functionalities of technology will have a good performance that impacts to achieve a task. Each technology has a distinct characteristic [18]. In this study, Leap motion can perform the task (drag and drop an object), but with poor performance [19].

C. Learning Outcomes

All participants who interacted with tangible-based interface and VR learner some scientific knowledge from the exhibit. Examples of knowledge obtained from the exhibit include scientific names, identification of edible mushroom , nutritional value of mushrooms, How to distinguish between good or bad mushrooms, different types of mushrooms, and frogs we can eat and cannot eat and touch.

In contrast, Only four people clearly stated that they learned some scientific content from the gesture-based interface. Others suggested they learned something from the interactive exhibit. Some common sentences about learning; ". . . don't eat the frog. . . ", ". . . different type of mushroom and their effect. . . ", and "...mushroom can harm me..".

A reason participants did not learn from the gesture-based exhibit might be the difficulty to pick up the object. They payed more attention to how to play with the interactive exhibit, how to move and pick up an object, and familiarize themselves with the technology. One participant said, "..it quite confuse...movement of my hand try to pick up the mushroom, I am not quite sure is it the right direction or not, . . . try to move my hand around the table. . . ". Another participant said, " the system is new for me. . . .I didn't pay attention to the instruction, because I try to familiarize myself to the system first..".

When designing an interactive exhibit for the STEM museum, the main aim is to deliver learning content. The results from this study indicate that pragmatic quality (efficiency, perspicuity, dependability) affect learning outcome. The difficulty and low efficiency of the system resist reaching the learning content of an interactive exhibit and reduces the chance to learn all content of the learning media.

D. System Feedback

After participants report they get some knowledge from the system, we investigate what feedback best delivers information to participants. The system design has multimodal feedback for the learner, which is a combination of three feedback modalities: visual animation, visual colour, and audio. Half of the participants reported that they learn from animation. Two participants reported on exhibit 2-tangible they learn from the coloured lights. Two participants reported on exhibit 3-VR they learned from audio. Other participants learned from a combination of two modalities; animation and colour, animation and sound, colour and sound. One participant stated that the colour (red, yellow, green) told them first and then the animation explained more in detail. The results show that learning occurs via all modalities of feedback, which corresponds to learners having various learning preference.

It has been found that bimodal feedback improves the learning experience and help the user continuously interact with the system effectively when other modality unable perform [20]. STEM museum environment can some time be loud from crowed visitors so using auditory feedback might not be effective in this situation. The feedback is most successful when occurring at the appropriate time, is related to the exhibit, and have meaning [21]. More consideration on the benefit of multimodal feedback of the system include observation that animation is effecive at explaining how the bio toxin harm human organs, sound and colour help the learner quickly indicate that it is dangerous mushroom frog.

However, Some participants confused the meaning of the feedback. The feedback system used an analogy technique to present information, with three different meanings. Some participants report that they were confused with the animation when the system showed yellow feedback. The participants actually interpreter the opposite meaning (dangerous/safe), but were unclear which feedback using analogy should consider the learner might misunderstand the content.

E. Recommendations for Design

Some of the participants suggested adding more information, the details about mushrooms and frogs. Some of the participants mentioned the quality of text must improve. Many participants stated the text information display in VR is not clear for them to read. Design exhibit for VR should consider the quality of text display in VR environment, it seems to deliver content via text inappropriate for learning through VR, especially using VR as an interactive exhibit in the museum. A VR headset used for various users with different sight, must allow for each user. In a museum settings , we cannot confirm the adjustment and fit to each user.

An advantage of the 3D object for STEM learning is perceived reality and detail of the object. From video recording, we found that many participants pay attention to the mushroom and frog models using either tangible-based interface or VR. A recommendation for the tangible-based interface is to make the model surface have more of a sense of touch a real texture. In VR, participant are able to hold the model and rotate it to see the detail on all sides. A recommendation for VR is to enhance the 3D model to be more realistic for learning.

F. Six aspects for choosing interface

To use the results from this study to explain or suggested a choice of technology between VR and alternative, specific design action for VR, we will discuss six aspects:

1) Novelty: it is an attractive feature for each interactive exhibit. The results show no difference between the interfaces. We can imply from the interview data that most of the participants have never seen in the museum or in their daily life before. The interfaces were quite new for them. If the interface is not usually used in their daily life, the user will feel that exhibit is new for them.
2) User-friendly: Tangible and VR technologies were found to same experience in this study. The majority of participants mentioned that easy to interact with the exhibits in both case. In contrast, the Gesture is quite difficult to play with and distracted player pay less attention to the content. Our results suggestion from choosing a technology should consider the ease of interacting with the system.
3) Precision of the input device: this factor is quite important. The input system with poor precision has a large effect on user experience. If the system is annoying to use, then they will stop playing. The problem of input precision comes from the precision of the input devices and the skill of the software developer.
4) Task and device design: design interaction for the system should consider the way people interact with the system suite when designing the input device. Action should be related to characters of input devices. Based on the experiment, some participants mentioned that the Leap motion device was inappropriate for use with the actions in this experiment (drag and drop the small object). It affects the system, makeing it difficult to use. One button and one action make people quickly understand and easy to interact with the system.
5) Multimodal feedback: the result from the experiment found that design feedback with multiple types of feedback do not make players confused, but each channel of the feedback (sound, colour, animation) help players better understand the content in a different ways.
6) Quality of text in VR: many participants mentioned that it was difficult to read the sentences in the VR environment, especially short sighted people. The use of VR as an exhibit should consider the quality of the text.

VI. Conclusion

The study chooses three different types of interface including Gesture-based interface, Tangible-based interface, and VR to develop three different interactive exhibits, but delivers the same scientific content and narrative style. The experiment has two hypotheses: H1-there is no difference between the user experience quality when using the gesture-based interface, tangible-based interface, and VR interface, and H2-VR has a higher holding power than the gestures-based interface and tangible-based interface.

The study uses the mixed-method quantitative and qualitative data gathering. The quantitative study measures user experience with a UEQ which summarize into six dimensions of user experience: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty [4]. The holding power is measured by the amount of time that players play with each exhibit. The qualitative study uses a semi-structured interview after the participant has an experience with each exhibit in order to gather user feedback in-depth detail and to find an issue of each exhibit. There are 31 participants in the experiment.

The statistical results were derived using the Friedman non-parametric test followed by a Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set by $\rho < 0.017(\alpha = 0.05)$ to analyse the difference between pair of technology. The results show that there is a significantly different experience when using different types of technology. There is an influence on user experience in terms of attractiveness, perspicuity, efficiency, dependability and stimulation dimension. There are differences between using UI01-Gesture and UI02-Tangible, and between using UI01-Gesture and UI03-VR. On the other hand, there is no effect on novelty dimension between using all pairs of technology in this experiment. There is not a statistically significant difference in playing time which indicates holding power of the interactive exhibit does not depend on the type of interactive exhibit interface.

Analysis of the results suggests six aspects of choosing an alternative technology to create a new interactive exhibit. They are the novelty, user friendly, precision of the input device, task and device design, multimodal of feedback, and quality of text in VR. The six aspects can be a guidelines for choosing technology, and comparing VR with other technologies to create an interactive exhibit for museum learning.

REFERENCES

[1] S. Bitgood, "Immersion Experiences in Museums,"in *Social Design in Museums, the psychology of visitor studies*, Edindurgh, UK: MuseumsEtc, 2011, pp. 102-17.

[2] H. Gilbert, "Immersive exhibitions: What's the big deal?," *Visitor Studies Today*, vol.5, no.3, pp.10-13,Sep.2001. Accessed on: Feb.2,2021.[online]. Available: http://kora.matrix.msu.edu/files/31/173/1F-AD-297-8-VSA-a0a6e0-a_5730.pdf

[3] C. Ramsay, "Hands-on,Hands-off: the personal, social and physical context of interactives in museums," in *ichim99*, 1999, pp. 27-36.Accessed on: Feb. 2, 2021. [Online]. Available: https://www.archimuse.com/publishing/ichim99/ramsay.pdf

[4] E. Hornecker and M. Stifter, "Learning from Interactive Museum Installations About Interaction Design for Public Settings", *OZCHI 2006*, Sydney, Australia, Nov. 20-24, 2006, pp. 135–142, doi: https://doi.org/10.1145/1228175.1228201

[5] J. Packer, "Learning for Fun: The Unique Contribution of Educational Leisure Experiences," in *Curator: The Museum Journal*, vol. 49, iss. 3, pp. 329-344, Jul. 2006, Accessed on: Feb. 9, 2021.[online]. Available: https://doi.org/10.1111/j.2151-6952.2006.tb00227.x

[6] P. Marshall, "Do tangible interfaces enhance learning," *TEI'07*, Baton Rouge, L.A., USA, Feb. 15-17, 2007, pp. 163-170

[7] Y. H. Wang, S. S. C. Young and J. R. Jang, "Evaluation of Tangible Learning Companion/Robot for English Language Learning," *2009 Ninth IEEE International Conference on Advanced Learning Technologies*, Riga, Latvia, 2009, pp. 322-326, doi: 10.1109/ICALT.2009.147.

[8] B. Ullmer and H. Ishii, "Emerging Frameworks for Tangible User Interfaces," in *Human-Computer Interaction in the New Millenium*, pp. 579-601, Aug. 2001. Accessed on: Feb. 6, 2021.[online]. Available: http://alumni.media.mit.edu/ ullmer/papers/tui-millenium-chapter.pdf

[9] NS. Chen and WC. Fang, "Gesture-Based Technologies for Enhancing Learning," in *The New Development of Technology Enhanced Learning* , Huang R., Kinshuk, Chen NS. ed., Berlin, Heidelberg: Springer, 2014, pp. 95-112 , doi: https://doi.org/10.1007/978-3-642-38291-8_6

[10] H. Sharp, Y. Rogers, and J. Preece,"Interface," *Interaction design beyond human-computer interaction*, 5th ed. Indianapolis, IN, USA:John Wiley&Sons, 2019, pp. 255-256

[11] H. Cecotti, Z. Day-Scott, L. Huisinga and L. Gordo-Pelaez, "Virtual Reality for Immersive Learning in Art History," *iLRN 2020*, San Luis Obispo, CA, USA, 2020, pp. 16-23, doi: 10.23919/iLRN47897.2020.9155108.

[12] K. Mahmoud et al."Does Immersive VR Increase Learning Gain When Compared to a Non-immersive VR Learning Experience?," in *Proc. 22nd HCII 2020 Conf.*, vol. 12206, Copenhagen, Denmark, July 19–24, 2020, pp 480-498, doi: https://doi.org/10.1007/978-3-030-50506-6_33

[13] D. Reinhardt and J. Hurtienne,"The Impact of Tangible Props on Gaming Performance and Experience in Gestural Interaction", in *Proc. TEI 2018 11th Int. Conf.*, Stockholm, Sweden, Mar. 18–21, 2018, pp.638–646, doi: https://doi.org/10.1145/3173225.3173258

[14] M. McEwan, A. Blackler, P. Wyeth, and D. Johnson,"Intuitive Interaction with Motion Controls in a Tennis Video Game," in *CHI PLAY 20 Proc. of Annu. Symposium on Computer-Human Interaction in Play*, Virtual Event, Canada, Nov. 2–4, 2020, pp. 321–333, doi: https://doi.org/10.1145/3410404.3414242

[15] M. Lykke and C. Jantzen, "User Experience Dimensions: A Systematic Approach to experiential Qualities for Evaluating Information Interaction in Museums", *CHIIR'16*, Carrboro ,N.C., USA, Mar. 13–17, 2026, pp. 81-90.

[16] S. Bitgood," Designing Effective Exhibit: Criteria for Success, Exhibit Design Approaches, and Research Stategies," *Visitor behavior*, vol.4, no.4, pp. 4-15, 1994. Accessed on: Feb. 2, 2021.[online]. Available: https://www.informalscience.org/sites/default/files/Designing_Effective_Exhibits.pdf

[17] Y. Georgiou, A. Ioannou, and M. Ioannou, "Investigating Children's Immersion in a High-Embodied Versus Low-Embodied Digital Learning Game in an Authentic Educational Setting," in *Proc. 5th Int. Conf. iLRN 2019*, London, UK, June 23–27, 2019, pp. 222–233, doi: https://doi.org/10.1007/978-3-030-23089-0_17

[18] R. Spies1 , S. Grobbelaar , and A. Botha, "A Scoping Review of the Application of the Task-Technology Fit Theory", in *Proc. 19th IFIP WG 6.11 Conf. on e-Business, e-Services, and e-Society, I3E 2020* , Skukuza, South Africa, Apr. 6–8, 2020, pp. 397–408, doi: https://doi.org/10.1007/978-3-030-44999-5_33

[19] L. E. Potter, J. Araullo, L. Carter, "The Leap Motion controller: A view on sign language," in *OzCHI '13: Proc. of the 25th Australian Computer-Human Interaction Conf.: Augmentation, Application, Innovation, Collaboration* ,Adelaide, Australia, Nov. 25-29, 2013, pp. 175–178

[20] H. S. VITENSE , J. A. JACKO, and V. K. EMERY, "Multimodal feedback: an assessment of performance and mental workload," in *Ergonomics*, vol. 46, no. 1-3, pp. 68-87, 2003, Accessed on: Feb. 9, 2021.[online]. Available doi: https://doi.org/10.1080/00140130303534

[21] B. Campbell and A. Feldmann,"THE POWER OF MULTIMODAL FEEDBACK," in *Journal of Curriculum, Teaching, Learning and Leadership in Education*, vol. 2 , iss. 2, pp. 1-6, Dec. 2017, Accessed on: Feb. 9, 2021.[online]. Available: https://digitalcommons.unomaha.edu/ctlle/vol2/iss2/1