A Comparison of Surface and Motion User-Defined Gestures for Mobile Augmented Reality

Author:

Ze Dong

Senior Supervisor:

Dr. Tham Piumsomboon

Co-Supervisor:

Prof. Rob Lindeman

Associate Supervisor:

Dr. Adrian Clark

Abstract

Augmented Reality (AR) technology permits interaction between the virtual and physical worlds. Recent advancements in mobile devices allow for a better mobile AR experience, and in turn, improving user adoption rate and increasing the number of mobile AR applications across a wide range of disciplines. Nevertheless, the majority of mobile AR applications, that we have surveyed, adopted surface gestures as the default interaction method for the AR experience and have not utilised three-dimensional (3D) spatial interaction, as supported by AR interfaces. This research investigates two types of gestures for interacting in mobile AR applications, surface gestures, which have been deployed by mainstream applications, and motion gestures, that take advantages of 3D movement of the handheld device. Our goal is to find out if there exists a gesture-based interaction suitable for handheld devices, that can utilise the 3D interaction of mobile AR applications.

We conducted two user studies, an elicitation study and a validation study. In the elicitation study, we elicited two sets of gestures, surface and motion, for mobile AR applications. We recruited twenty-one participants to perform twelve common mobile AR tasks, which yielded a total of five-hundred and four gestures. We classified and illustrated the two sets of gestures, and compared them in terms of goodness, ease of use, and engagement. The elicitation process yielded two separate sets of user-defined gestures; legacy surface gestures, which were familiar and easy to use by the participants, and motion gestures, which found to be more engaging. From the design patterns of the motion gestures, we proposed a novel interaction technique for mobile AR called TMR (Touch-Move-Release). To validate our elicited gestures in an actual application, we conducted a second study. We have developed a mobile AR game similar to Pokémon GO and implemented the selected gestures from the elicitation study. The study was conducted with ten participants, and we found that the motion gesture could provide more engagement and better game experience.

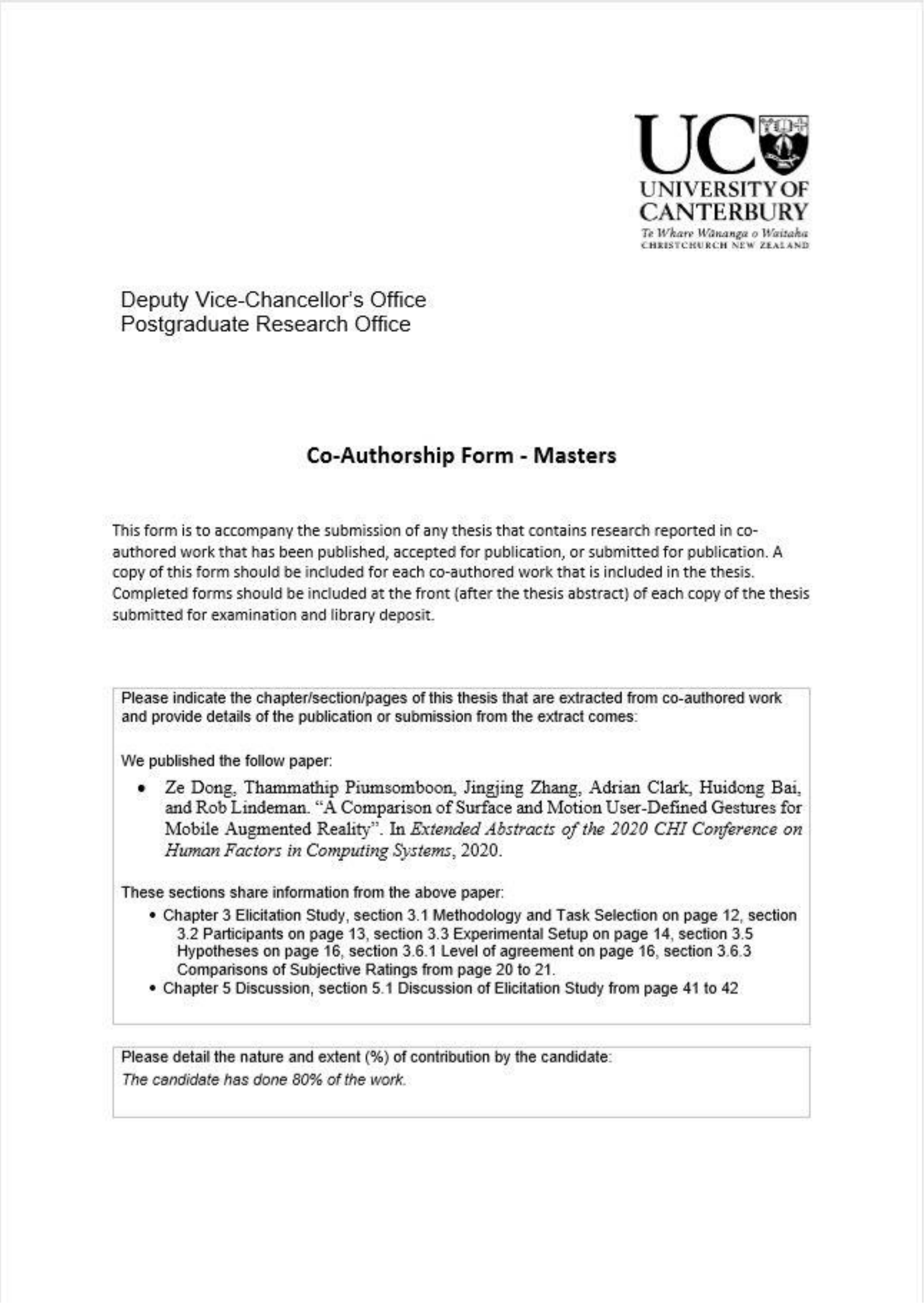
Nevertheless, surface gestures were more accurate and easier to use. We discussed the implications of our findings and gave our design recommendations for designers on the usage of the elicited gestures. Our research can be further explored in the future. It can be used as a "prequel" to the design of better gesture-based interaction technique for different tasks in various mobile AR applications.

Keywords: augmented reality, mobile device, gestures, elicitation study

Publication

A peer-reviewed publication listed below is based on research completed as part of this thesis:

- Ze Dong, Thammathip Piumsomboon, Jingjing Zhang, Adrian Clark, Huidong Bai, and Rob Lindeman. “A Comparison of Surface and Motion User-Defined Gestures for Mobile Augmented Reality”. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.



UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Deputy Vice-Chancellor's Office
Postgraduate Research Office

Co-Authorship Form - Masters

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

We published the follow paper:

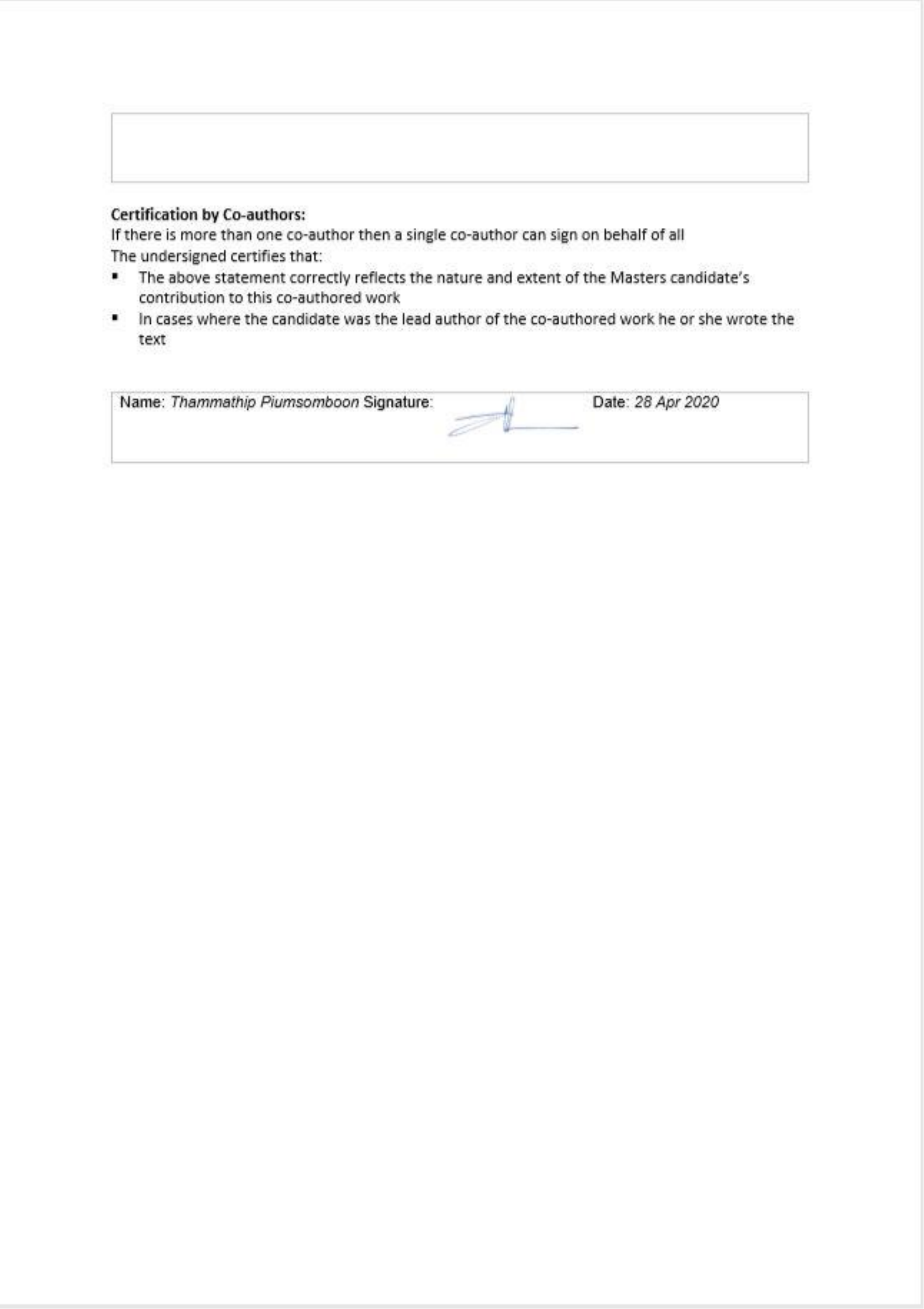
- Ze Dong, Thammathip Piumsomboon, Jingjing Zhang, Adrian Clark, Huidong Bai, and Rob Lindeman. "A Comparison of Surface and Motion User-Defined Gestures for Mobile Augmented Reality". In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

These sections share information from the above paper:

- Chapter 3 Elicitation Study, section 3.1 Methodology and Task Selection on page 12, section 3.2 Participants on page 13, section 3.3 Experimental Setup on page 14, section 3.5 Hypotheses on page 16, section 3.6.1 Level of agreement on page 16, section 3.6.3 Comparisons of Subjective Ratings from page 20 to 21.
- Chapter 5 Discussion, section 5.1 Discussion of Elicitation Study from page 41 to 42

Please detail the nature and extent (%) of contribution by the candidate:

The candidate has done 80% of the work.



Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifies that:

- The above statement correctly reflects the nature and extent of the Masters candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: *Thammathip Piumsomboon* Signature: Date: *28 Apr 2020*

Acknowledgement

In the process of completing this thesis, I would like to thank my senior supervisor, Dr Tham Piumsomboon, co-supervisor, Prof. Rob Lindeman, and associate supervisor, Dr Adrian Clark, for their advice and guidance. I feel honoured to have them as my supervisors. I would also like to thank the examiners of this thesis for taking the time to read and evaluate my work. I would like to grateful for the assistance from HIT Lab NZ staff and students. Also, I am appreciative of my family for their support, especially my wife, Zhang Jingjing, who helped me through the most difficult time.

Table of contents

Chapter 1

Introduction

According to Azuma [2], Augmented Reality (AR) can be achieved if the AR systems are capable of demonstrating three characteristics. Firstly, they must be able to combine real and virtual imagery. Secondly, they have to support real-time interaction. Lastly, they must be able to register virtual content in 3D space. At the time of this writing, arguably the most advanced commercial AR system available would be the Microsoft HoloLens 2 [14], a head-worn AR device that is capable of achieving all three characteristics. Nevertheless, HoloLens 2 is still out of reach of the general public and aim at professional use cases. On the contrary, handheld devices, such as mobile phones, are ubiquitous, and they are currently the primary way for people to experience AR in a variety of domains [7, 55]. For this reason, it is our motivation to explore user interaction, focusing on mobile AR systems through mobile devices with touchscreen support.

Recent advancements in mobile technology have led to an increasingly wide range of mobile applications which use AR as their core mechanic for visualisation and interaction. Mobile AR enabling frameworks, such as Apple's ARKit [30] and Google's ARCore [31], have made the development of mobile AR applications accessible to more developers than ever. This leads to a variety of AR applications in several domains. For example, IKEA Place [3] allows customers to visualise virtual furniture in their home (see Figure 1). QuiverVision [47] is the first to introduce AR colouring books, and SketchAR [51] teaches users how to draw by overlaying virtual drawings over a real canvas. From a survey, we have found that most mobile AR applications have adopted existing interaction metaphors based on surface gestures designed for devices with a touch-sensitive screen. The touch input is implemented as part of the underlying input of the mobile platform framework, which is a dominant and familiar method of interaction for regular mobile users. Nonetheless, past research has demonstrated methods beyond those currently used in the mobile AR applications to enrich mobile AR experiences.

Surface gestures are the conventional interaction technique used in handheld mobile devices and adopted by mobile AR applications. Previous studies explored various design principles of surface gestures using different methodologies ranging from expert's design [59, 61], participatory design by non-experts [60], or comparative studies of both groups [37]. Nevertheless, surface gestures have their drawbacks in some way. For example, the handheld devices only have limited interaction area for surface gestures [6], which is also restricted to 2D [4, 28], and only a limited number of fingers are able to fit in such area [19]. Furthermore, gesturing on the screen tends to cause occlusion

[18] and focusing on the on-screen interaction may lead to dual perspectives [13]. Another type of gestures, mid-air gestures, are widely used with the AR head-mounted display (HMD), as they can offer 3D interaction [41]. However, mid-air gestures are not ideal to use in public [48], prolong usage can lead to fatigue [26], and bimanual gestures are not possible on the handheld mobile device. Further investigation has led us to the third type of gestures, the motion gestures, which utilise the mobile device built-in sensors to detect the device movements. Past research has proposed and demonstrated motion gestures as the interaction technique for handheld devices [1, 25, 27, 32, 49] or in conjunction with a secondary device [9, 52]. There have been examples of motion gestures used in the AR context, for example, direct camera manipulation [25], or virtual object manipulation [22, 38]. However, to our knowledge, there has not been any research that explores the participatory design of motion gestures for a broad range of mobile AR applications nor compares them against the conventional interaction techniques to validate their usability.

Figure 1: IKEA Place user place an AR armchair on the deck

Due to the limited knowledge available on motion gestures for mobile AR interaction, we have decided to conduct a study to explore further gestures suitable for different tasks in the mobile AR context. We have chosen to pursue a participatory design methodology, specifically, an elicitation study [12]. We have adopted the method of Wobbrock et al. [60]. There have been elicitation studies conducted for motion gestures for handheld devices [49], and gestures for AR context [43] in the past. However, the former study focused on eye-free interaction in a non-AR context, while the latter emphasised on gestures for the head-worn AR system. Our goal is to explore the gesture design space for handheld devices and find out if there exists a gesture-based interaction, that can utilise the 3D space that mobile AR applications support. For our elicitation study, we have conducted a survey of common tasks among popular mobile AR applications. We have selected twelve gestures to elicit two sets of gestures, surface and motion gestures, from the participants, and to compare their subjective ratings. We hypothesis that the surface gestures set would be rated higher in terms of suitability and ease of use, while motion gestures would be more engaging. In a follow-up study to validate our gestures, we have implemented a Pokémon GO Clone, a mobile AR game, and implemented the two elicited gestures for a throwing task. We hypothesise that there would be differences in terms of accuracy, subjective ratings, in-game experience, and system usability between the two interaction techniques. This research has contributed to the following outcomes as a result:

1. A literature review in the areas of surface, mid-air, and motion gestures, mobile AR interaction, and previous research with elicitation studies.
2. An elicitation study yielded two sets of user-defined surface and motion gestures, anecdotal feedback, and the results of a comparison between the two gestures in terms of suitability, ease of use, and engagement.
3. An overview of the development of a mobile AR application, a Pokémon GO Clone game, and the implementation of the selected surface and motion gesture elicited from the previous study.
4. A validation study compared the chosen gestures in an actual mobile AR game, examining the two gestures in terms of accuracy based on three levels of target sizes, subjective ratings from the previous study, in-game experience questionnaire, system usability scale, and user preferences.
5. From the results of both studies, we have summarised and discussed our findings and provided their implications and guidelines. We have proposed the TMR (Touch-Move-Release) interaction technique for mobile AR applications.

In the chapters to follow, we cover our literature review in Chapter 2. Chapter 3 reports the results of the elicitation study. Chapter 4 provides an overview of the mobile AR application development and gestures implementation, as well as the experimental details and results of the validation study. Our discussion of the outcomes of both studies and their implication will be covered in Chapter 5. The conclusion and future work are presented in Chapter 6.

Chapter 2

Related Work

In this chapter, we cover the background research of related topics into gesture-based interaction techniques as well as the methodology for participatory design specifically, elicitation studies. Previous research has been categorised into four subsections. We provide a brief overview of research on surface and mid-air gestures in Section 2.1. We introduce the interaction technique of interest in motion gestures in Section 2.2. We cover past mobile AR interaction and state of the art in Section 2.3. The previous elicitation studies will be discussed in Section 2.4. Finally, Section 2.5 covers our research questions and goals.

2.1 Surface and Mid-Air Gestures

Surface gestures have been a fundamental method of interaction for surface computing, which utilises a touch-sensitive screen as the primary input. Previous research provided various guidelines for designing and implementing surface gestures. Wobbrock et al. [60] proposed a taxonomy and user-defined surface gestures set from twenty participants who were regular people without training in the area of interaction design and provided guidelines and implications of their findings. In the follow-up study [37], they compared the elicited user-defined gestures set to the elicited experts' set created by three interaction design experts. They found that gestures developed by the majority of users were rated higher, which was also true for the gestures proposed by multiple researchers. Furthermore, although some of the researcher's gestures were found attractive, but the participants ultimately chose simpler gestures that took less effort to use. Wu et al. [61] proposed three design aspects of gesture registration, gesture relaxation, and gesture and tool reuse, which considered the interaction context, comfort level, and applicability of each gesture to different tasks, respectively. For a preliminary study, they developed a prototype application for a tabletop surface computing system and implemented four types of gestures, including annotate, wipe, cut/copy-n-paste, and pile-n-browse. In another approach, Wilson et al. [59] focused on improving the realism of surface interaction through physics simulation by creating proxy particles to exert force on the virtual objects. They conducted an experiment with six participants to complete three physics-based tasks of positioning, sorting and steering. They found that interaction through the proxy particles could shorten task completion time and received positive feedback for the proposed technique. Nevertheless, the challenges exist when using surface gestures on a small screen of the handheld device such as limited interaction space unable to support some gestures [4], hand occlusion of the display [18], limited reach when operating single-handedly [6]. Further discussion on surface gestures in mobile AR context will be covered in Section 2.3.

Mid-air gestures or gesturing in-the-air, refers to gestures which are performed while holding an arm or arms in front of one's body such as pointing, pushing, waving etc. Through the physical nature of the arms and hands' movement, these gestures could provide an immersive experience while interacting. For instance, Cui et al. [15] investigated the user's mental models while performing mid-air gestures for shape modelling and virtual assembly with sixteen participants. They found that users had different preferences for interaction technique and felt more natural and comfortable for their preferred method, and bimanual gestures were more natural than unimanual. Kyriazakos et al. [34] designed a novel fingertip algorithm to extend the interaction method. The interaction between the user and the virtual object was achieved by tracking mid-air gestures using the rear camera of the mobile device. For example, the user could use a victory hand pose to move the virtual object using the two fingers. Previous studies also have applied mid-air gestures in various settings, such as on public display [57], pairing between an armband sensor and a handheld device [16]. Although mid-air gestures could take advantages of performing in the 3D space, using them for an extended period could also lead to fatigue and discomfort. Rico and Brewster [48] also raised an issue around the social acceptability of mid-air gestures performing in public. In their study, they had the participants watched demonstration videos of different gestures and asked them to imagine performing those gestures in a different social environment. They found that social environment did affect the use of gestures, and device-based gestures were more socially acceptable. To validate their findings, they had eleven participants performed chosen gestures in public and found that gestures that attracted less attention were more preferable by the participants. They recommended that the designer must avoid emblematic gestures, which might lead to confusion of context and use more familiar and socially acceptable gestures, especially to be used in public.

Surface gestures are a common method for users to interact with mobile devices and are strictly 2D in nature. On the contrary, mid-air gestures are performed in 3D space, bringing flexibility and more possibilities to the design space restricted by the recognition technology. Previous research has shown that two-handed mid-air gestures are commonly elicited especially in AR tasks. However, bimanual gestures are not possible on the mobile devices as the users are required to hold the device in one hand. Past research had also explored another type of gestures which utilised the movement of the handheld device as inputs. These gestures are known as motion gestures. We will be covered these gestures in the next section.

2.2 Motion Gestures on Handheld Devices

Motion gestures are another type of interaction method on handheld devices that utilise inertial measurement unit (IMU), which combined an accelerometer and a gyroscope to obtain the orientation and linear acceleration to track the device's movement. Previous research investigated motion gesture as an alternative input method for handheld devices. Motion gestures typically involve more hand or arm's movement during the interaction providing unique experiences complementing those of surface interaction methods. Hinckley et al. [27] integrated multiple sensors, including proximity range, touch sensitivity, and tilt sensors and introduced novel functionalities on a handheld device. For example, the device would wake up when it was picked, and scrolling could be achieved by tilting the device. They found that sensors opened up new possibilities allowing a vast interaction design space for the handheld devices.

Wigdor and Balakrishnan [58] proposed TiltText, an interaction technique that could reduce ambiguities of the text input process using a combination of a keypad and four tilting directions, left, right, forward, and back. They found that TileText was faster to perform, despite the higher error rate. Hartmann et al. [23] explored the role of sensors in the handheld interaction, dividing the development process into three simple steps of connecting the appropriate hardware, creating the logic, and establishing the relationship between sensors and logic. They proposed a tool to help interaction designers to map the connections between the sensors and applications to support direct manipulation and pattern recognition. They showed that sensors had become a crucial tool for the interaction designer to enhance the interaction and overall user experience.

GripSense system [19] explored one of the possible solutions to address the limitation of a single-handed surface interaction such as the challenge to replicate the "pinch-to-zoom" gesture with one hand. They made use of the touch input in conjunction with inertial sensor and vibration motor to detect the level of pressure exerted by the users on the screen. The technique allowed complex operations to be performed with a single hand. Ashbrook et al. [1] raised concerns regarding motion gesture design emphasising two points. Firstly, the gestures proposed by designers might not be practical for the actual recognition technology. Secondly, how the system could ensure the robustness of the recognition system to avoid false registration or activation. To address these problems, they proposed MAGIC, a motion gesture framework that defines the design process in three stages, requirements gathering, determine the function activation, and user testing. This process enabled non-experts to leverage sensors, e.g. accelerometers, in the design process of motion gestures. GesText [32] was another system that made use of an accelerometer to detect motion gestures for text inputs. They found that the area-based layout supported by a simple tilt motion gestures was more efficient and preferred over the alphabetical layout.

In another application, motion gestures could be performed on a handheld device to provide inputs and interaction with the virtual objects on a large display system [8]. In terms of human-centred design, Ruiz et al. [49] conducted an elicitation study with twenty participants to collect motion gestures for nineteen tasks on a handheld device. They categorised the gesture dataset based on two aspects, gesture mapping and physical characteristics. They found that the mapping of commands influenced the motion gestures consensus. Past research had demonstrated that the handheld device's sensors could be leveraged for the recognition of motion gestures through the movement of the device. The benefits of motion gestures inspired us to further explore the design space in the context of mobile AR application. We propose a comparison between surface and motion gestures on a handheld device for mobile AR, which we believe has not been investigated prior to this research.

2.3 Mobile AR Interaction

To date, there had been a number of researches proposing various methods of interaction in mobile AR, which offered different experiences for user interaction. Early research demonstrated that mobile AR could provide precise 6-DOF (degree-of-freedom) camera/viewpoint control through the movement of the handheld device using the device registration in the physical environment. In one of the first face-to-face collaborative mobile AR application, Henrysson et al. [25] developed AR Tennis with two users sitting across the table and played a game of virtual tennis on the table. The mobile device's rear camera registered an image marker placed on the table to determine the device 6DOF location relative to the marker. To interact with the virtual tennis ball, the user could move the camera in front of the ball's incoming path and nudge the device forward to exert force onto the ball in order to hit it back. Through the user study and feedback, they provided guidelines for designing games for mobile AR such as do provide multi-sensory feedback, focusing on the interaction, and support physical manipulation. They found that the combination of visual, tactile, and auditory outputs during the interaction process and the camera manipulation offered by AR systems further increased the level of immersion and could improve collaboration and entertainment.

The AR-Tennis prototype inspired Ha and Woo [22] to develop the ARWand, a 3-DOF device for mobile AR interaction based on the device's sensor. The users could use surface gestures on the device's screen to manipulate virtual objects for example when the device was held vertically to the ground, swiping up or down would move the object higher or lower, and when it was held horizontally, the swipes would manipulate the forward and backward direction instead. This technique supported an individual axis control; however, relying on the built-in sensors alone might not yield best in terms of precision. Later, Mossel et al. [38] proposed HOMER-S, a 6-DOF interaction technique to support object manipulation. The user could perform translation or rotation via the device's touch interface. The surface gesture could be performed on the object’s gizmo to change the object’s position and orientation. This technique supported single-handed operation and could complete the task faster at the cost of lower accuracy.

To address the issue of shaky hand and improve the accuracy of the manipulation, Lee et al. [35] proposed a technique called "Freeze-Set-Go", which allowed the user to pause the current viewpoint of the AR system to let the user to manipulating the object in the current view. Once the task was completed, the user could resume the normal tracking of the viewpoint and update the virtual object’s location accordingly. They found that this technique helped improved accuracy and reducing fatigue. Tanikawa et al.[53] created a mobile AR system to support multimodal inputs of viewpoint, gestures, and device's movement. They demonstrated the interaction in the AR Jenga game where the user could touch the screen to select the virtual wooden block to move. While holding the finger on the screen, the block was kept at a fixed distance from the screen. The user could move the block by moving the device, and when the finger was removed from the screen, the block could be released. They found that this technique provided reasonable accuracy and offered a better experience for object manipulation in mobile AR. Beyond improving the accuracy of the manipulation technique in mobile AR, the dual perspectives problem [13] was another issue that impacted the mobile AR interaction. This was when the viewpoint captured by the device’s rear camera and displayed on the screen would not match the scale of the world viewing through the user’s actual perspective. Furthermore, ergonomic was also identified as another limitation of handheld-based mobile AR. Colley et al. [11] evaluated the ergonomics of the camera placement of mobile AR devices by comparing the level of tilt of the camera to the screen. They found that the screen size and a proper tilt level had a significant impact on the level of comfort while interacting.

In another approach to enhance mobile AR interaction, researchers explored either combining multiple devices for multi-device inputs or using mobile AR system for visualising embedded sensors. Goldsmith et al. [21] demonstrated SensAR, combining a mobile device and environmental detection sensors. When a marker was scanned using their handheld device, the detection sensor shared the environment data and displayed them in AR. They found that the user found this method of interaction and visualisation to be seamless and immersed. Stanimirovic and Kurz [52] took this further and used a smartwatch’s camera to scan markers for hidden AR content scattered in the environment so the user could use their handheld device to access them. Chen et al. [9] used a smartwatch to edit a clipboard text and sent the update to the handheld device. Nevertheless, multi-device interaction required additional devices to operate and might not be ideal for mobile AR applications in general.

Despite significant advancements in mobile AR technology, challenges exist in the development of better mobile AR interaction and overall experience [33]. Hürst and Van Wezel [28] found that previous mobile AR interaction was limited to the 2D screen interaction, such as touching and swiping, introducing an issue of screen occlusion that could hinder the operation. Instead, they proposed a technique to use the device’s rear camera to track the user’s thumb and index finger for direct manipulation of the virtual object. They found that this approach could offer more natural and did not occlude the screen. However, the user's fingers must remain in the view of the camera, and both hands would be required to manipulate the object. Similarly, Bai [4] proposed mobile AR interaction behind the handheld device. He evaluated the mobile AR interaction method on the existing platform. He found that 3D gesture-based manipulation was more intuitive, engaging, and utilising 3D nature of AR than using surface gestures and could be less fatiguing than motion gestures. Through the literature review of the past mobile AR interaction, we found that interaction based on the device’s movement could offer the users 3D interaction and physical exertion [42]. However, accuracy was also an issue in 3D interaction. We observed that interaction techniques, which combined touch input and device movement, could improve the precision of the interaction utilising the touchscreen while preserving 3D interaction experience. For this reason, we decided to investigate motion gestures, which combined a touch input and device movement in this research.

2.4 Elicitation Study

Elicitation research utilises a method of collecting knowledge by analysing the behaviour patterns and feedback data of participants [12]. For example, Voida et al. [56] explored interaction with projection display in an office environment. They mapped the user's mental models by observing the user's manipulation of 2D objects in an AR environment. They asked the participants to propose gestures to interact with multiple projection displays and found that the pointing gesture was commonly used from afar. But when the virtual objects were closer to the participants, a user interface was preferable. Epps et al. [17] studied the user preferences of the tabletop interaction through an elicitation study where they displayed images depicting different tasks on the desktop and asked users to come up with gestures that they would use to perform the tasks. They presented gesture guidelines for the hand poses and corresponding tasks for tabletop systems from twenty participants. They found that over seventy percent of the time, the index finger was frequently used in multiple tasks, such as tapping, drawing, or swiping.

Elicitation studies were also used to create new taxonomy and collection of user-defined gesture sets. For surface computing, Wobbrock et al. [60] conducted a study that allowed non-expert participants to design surface gestures and evaluated the quality of those gestures in terms of suitability and ease of use on a tabletop system. With one thousand and eighty gestures elicited from twenty participants, they proposed a taxonomy and user-defined surface gestures set based on the gestures with the consensus score. With a think-aloud protocol, they could understand the users' design process, thereby provided guidelines for designers. Ruiz et al. [49] applied the same elicitation procedures for motion gestures performing on a handheld device. The elicited motion gestures exhibited characteristics across two dimensions of movement and mappings of commands. They provided anecdotal findings to help guide designers to design better gestures that would mimic everyday tasks and how sensors could be used to recognise those motion gestures. Later, Piumsomboon et al. [43] adopted the methodology and elicited gestures for an AR head-mounted display system. They extended Wobbrock 's taxonomy and proposed forty-four user-defined gestures for AR. They found that the majority of the gestures were performed mid-air and most of them where physical gestures that mimicked direct manipulation of the objects in the real world. They also found that similar gestures shared the same directionality with variants of hand poses. The anecdotal findings and implications were provided to guide designers in the their gesture design for AR.

Past research had also conducted a comparative elicitation study. Hayati et al. [24] investigated two interaction techniques, on-skin and freehand gestures, as they did not require an intermediate device for input. They compared two user-defined gesture sets on four aspects, social acceptability, learnability, memorability, and suitability. With a total of twenty participants, they found that the on-skin gestures with the small movement were better for social acceptance, while participants found freehand gestures to be better for immersion. Chen et al. [10] extended the research to explore inputs on the other body parts. They also collected the user-defined gesture sets and validated them with another group of participants. Their method departed from the previous elicitation studies. They combined the subjective score and physiological risk score. They found that gestures combined with the body parts helped enhance the naturalness of the interaction. There were also elicitation studies which compared gestures under different use-cases and scenarios. May et al. [36] elicited mid-air gestures to be used specifically within an automobile. They found that a participatory design process yielded gestures which were easier to understand and to use than the ones designed by the designers. Tran et al. [40] also elicited mid-air gestures in three different settings, mid-air, surface, and room, for varying virtual object's sizes. It was found that the scale of the target objects and the scenes influenced the proposed gestures.

The elicitation studies yielded user-defined gestures that were found to be simple and easy to use, allowing the designers to observe the design pattern for various constraints and settings and reflected the user's mental model under the given circumstance. Previous studies had shared design guidelines for various gestures, whether it be surface, motion, or mid-air gestures for different tasks, systems, and scenario. Nevertheless, we have not encountered any research that has elicited gestures for mobile AR settings. We still have limited knowledge of the design practice and guidelines of motion gestures for mobile AR interaction. Moreover, there has not been any comparison of the performance between the surface and motion gestures for mobile AR interaction. To address this shortcoming, we have conducted an elicitation study to elicit the surface and motion gestures for mobile AR to understand the possibilities in the proposed design space.

2.5 Research Questions and Goals

Our research interest lies in the potential enhancement of the user experience of mobile AR applications by exploring novel interaction techniques. Through the literature review, we discovered common mobile AR interaction methods, which were surface gestures, mid-air gestures, and motion gestures. The touchscreen-based surface gestures were the most widely used method on handheld devices and were highly familiar interaction for regular mobile users. However, the interaction was limited to the 2D screen and did not support nor able to utilise the immersive 3D experience of mobile AR. It was also found to offer limited interaction area [6] and suffered from screen occlusion [18]. We also found that mid-air gestures could not be used at their full potential due to the one hand would be holding the handheld devices as opposed to the AR head-mounted display systems where the users could operate using both hands. Therefore, we have decided not to investigate mid-air gestures. Instead, we are exploring a motion-gesture-based interaction technique which combines a touch input and a device movement to provide 3D interaction for mobile AR. To our knowledge, the proposed interaction technique has not been well explored, and so we have limited information regarding the performance of the surface gestures and motion gestures for the mobile AR setting. Through the investigation into existing mobile AR applications, we have identified the common tasks suitable for the elicitation process and comparison of the two types of gestures. This research gap raised a number of research questions:

RQ1 – Is there gesture-based interaction that can utilise three-dimensional space that mobile AR applications support for a handheld device?

RQ2 – How would the users perceive the proposed set of gestures compare to conventional surface gestures in terms of suitability, ease of use, and engagement?

RQ3 – How do these gestures interaction technique fair against surface gestures in an actual mobile AR application?

To answer these research questions, we have conducted an elicitation study for both surface and motion gestures, which we will cover in Chapter 3. We have also validated our results by selecting one of the tasks and comparing the gestures based on the actual implementation in a working mobile AR game, and this will be presented in Chapter 4. We have summarised our findings and discuss our results in Chapter 5. We conclude our research outcomes and our plan for future work in Chapter 6. We believe that the outcomes of this research will help designers to understand better the benefits and drawbacks of surface gestures and motion gestures for mobile AR interaction.

Chapter 3

Elicitation study

To answer our first research question, “Is there gesture-based interaction that can utilise three-dimensional space that mobile AR applications support for a handheld device?”, we have to find gestures that are best suited for various tasks in mobile AR applications. We have conducted an elicitation study for the insights from the users themselves. Our goal is to discover potential gestures that are able to utilise the 3D interaction space of the mobile AR applications running on a handheld device. We call these gestures, motion gestures, similar to the work by Ruiz et al. [49]. However, our definition is broader and does not limit these gestures to just the movement of the handheld device in 3D space but also consider interaction which combines any device’s movement and touch inputs from the device’s touchscreen.

Ruiz et al. proposed that these motion gestures could be recognised using built-in sensors of the hand-held devices. Current AR technology combines both software and hardware techniques, computer vision and sensor fusion, to localise the device’s 6 DOF (degree-of-freedom), position and orientation, in the physical environment. By incorporating the device’s 6DOF manipulation and the touch inputs of the handheld devices with mobile AR capability give rise to potentially novel motion gestures that have not been explored. We believe that mobile AR applications should be able to take advantage of such motion gestures with the combination of the handheld device’s movement and touch inputs. Nevertheless, surface gestures have been the dominant form of interaction for handheld devices with touchscreen input. Therefore, it is crucial to identify the benefits and drawbacks of these two types of gestures when using in mobile AR context. This comparison would provide us with answers to our second research question, “How would the users perceive the proposed set of gestures compare to conventional surface gestures in terms of suitability, ease of use, and engagement?”.

In this chapter, we share the results of our elicitation study to help answer our first two research question, RQ1 and RQ2. We discuss our methodology and task selection in Section 3.1. We provide details of participants, experimental setup and the procedure in Section 3.2, 3.3, and 3.4, respectively. In Section 3.5, we propose our hypotheses, report the results of the study in Section 3.6, and finally give a brief summary in Section 3.7.

3.1 Methodology and Task Selection

We have adopted an elicitation technique proposed by Wobbrock et al. [60]. This method selects common or interesting tasks that the targeted system should support. The researchers would first develop a set of descriptions of the tasks or short animations or videos, which depict the effects of the manipulation. This removes the need to develop a gesture recogniser circumventing the limitation of the underlying technology and remove any constraint in the design process. During the elicitation study, the participant would be informed of the task that they design the gesture for. The animations or videos of the selected task would be displayed to the participant using the system's display, e.g. a large surface computing touchscreen [60], AR headset [43], or simply the descriptions of the task [49]. Once the participant comprehends the effect of the task, they would be asked for its cause, i.e. what might be the gesture that yields the given outcome. After eliciting the gesture, the participant has to rate their gesture for the goodness-of-fit (*Goodness*) and ease-of-use (*Ease of Use*) on a 7-point Likert scale. In our study, we asked our participants to watch two videos and designed two gestures, a surface gesture and a motion gesture. Furthermore, apart from the goodness-of-fit and ease-of-use, we introduced the third measure and asked our participants to rate their gestures in terms of engagement (*Engagement*) as well.

To come up with appropriate tasks for the elicitation study, we surveyed sixteen mobile AR applications on both Google Play and the Apple App Store. In the end, we selected twelve tasks from six mobile AR applications as shown listed in Table 1, based on their high level of commonality across applications, and those appeared in the past research [20, 43, 49, 60]. To elicit the gestures, we prepared a set of videos for the twelve tasks by recording the screen during the interaction from the six chosen mobile AR applications. Two videos were recorded for each task. The first was for the surface gesture mimicking the manipulation of the virtual object with minimal movement of the mobile device. The second video was for the motion gesture with some movement of the mobile device during the manipulation mimicking the virtual object movement. From the two videos, we collected two sets of user-defined gestures for the surface gestures and the motion gestures for different mobile AR tasks.

3.2 Participants

For this research, we applied for approval through an ethics application reference number of HEC 2019/94/LR and the application was approved by the Human Ethics Committee, the University of Canterbury on the 14th of November 2019. The participants had to sign a consent form, which contained the experiment information. The participants were told that they could discontinue the experiments at any time without penalty regardless of the fact that there was not any serious health and safety concern. We provided the participants with a gift voucher for their participation in the study. Twenty-one participants (ten females) were recruited, aged 18 years to 59 years old, with an average age of 29 (SD=10.7) years. They were all right-handed. All of the participants owned a touch-screen mobile device, 8 had no prior experience with any mobile AR application, and the remainder had some experience, but none were frequent users.

Table 1: A list of selected tasks from six common mobile AR applications.

#	Apps	Tasks	#	Apps	Tasks
1	Angry Birds	Slingshot	7	IKEA Place	Move
2	Pokémon GO	Throw	8	IKEA Place	Rotate
3	Just a Line	Draw	9	ARia	Open Drawer
4	Just a Line	Erase	10	ARia	Close Drawer
5	CooolAR	Scale Up	11	CooolAR	Open Door
6	CooolAR	Scale Down	12	CooolAR	Close Door

3.3 Experimental Setup

The setup for this study was kept simple; the participants were seated in front of a television screen, while the experimenter was seated to the right of the participant, as shown in Figure 2. The participants were given a mobile phone to hold, a Samsung Galaxy S9, as a prop in the design process. To overcome the limited screen size of the mobile phone and the finger occlusion issue during the gesture design process, we chose to display the recorded videos on the 32” television screen placed in front of the participants instead. This way, the participants could watch the video and perform the gesture on the mobile phone at the same time. The participants were asked to follow a think-aloud protocol, and their gestures were recorded with a camera rig set up behind them to the right-hand side as they were all right-handed.

3.4 Procedure

We describe the procedure of the study in greater details as follows:

a. The experimenter introduced themself and made safety recommendations to the participants. Participants were informed that they could stop the study at any time.

b. Participants were asked to read the information sheet and sign the consent form (see Appendix A.1 and A.2). The experimenter then explained or clarified any question or concern that the participants might have in more detail. The experimenter confirmed that the participants gave their consent for the video recording of their interaction during the experiment.

c. Participants were informed of the procedure of the elicitation study. A pre-experimental questionnaire (see Appendix A.3) was presented to the participants for collecting the basic information and their previous experience with mobile AR applications if any.

d. Before starting the elicitation process, we gave the participants two minutes to familiarise themselves with the setup and allowed them to ask any question. As the process began, the participants were asked to watch the video and design a gesture for the task illustrated in the video.

e. After the gesture was elicited for each task, the participants were asked to rate their gestures on a 7-point Likert scale in terms of *Goodness* (how suitable was it for the task?), *Ease of Use* (how easy was it to perform?), and *Engagement* (how engaging was it to use?). Each task took approximately 4 minutes, for eliciting the two gestures, surface and motion.

f. Finally, after completing the elicitation process, we presented a post-experiment questionnaire (see Appendix A.4) to the participants for collecting their general feedback for the experiment. The study took approximately an hour to complete.

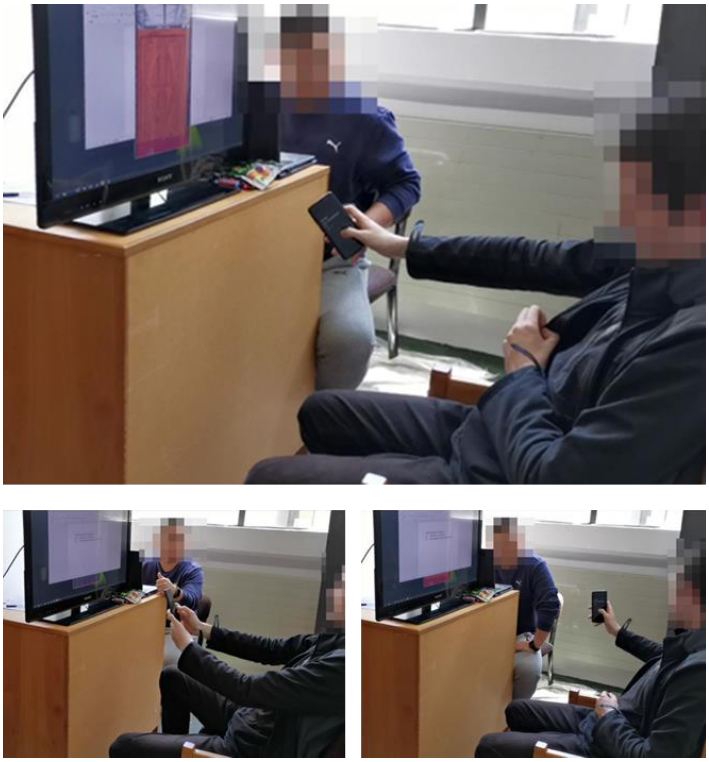


Figure 2: Experimental Setup - a participant is performing a gesture while watching a video displayed on a TV screen.

3.5 Hypotheses

For the surface gestures, which were well established, universal, and highly familiar to regular users of mobile phones and tablets, we expected to elicit a legacy set of common surface gestures being used in mobile AR applications. As a result, we hypothesise that ***the participants would rate the surface gestures higher in terms of Goodness and Ease of Use than the motion gestures (H1)***. Nevertheless, we believed that the movement required to perform motion gestures would enhance the user experience. Therefore, ***the participants would find the motion gestures more engaging than the surface gestures (H2)***.

3.6 Result

The elicitation study yielded two sets of user-defined gestures, surface and motion gestures, for a total of 504 gestures from twenty-one participants watching twenty-four videos for the twelve selected tasks. We illustrate the number of common surface and motion gestures elicited for each task in Figure 3. We report the level of agreement and their characteristics of the elicited gestures in subsection 3.6.1 and 3.6.2, respectively. We classified and illustrated the two sets of gestures, and compared their subjective ratings in terms of *Goodness*, *Ease of Use*, and *Engagement* in subsection 3.6.3. In subsection 3.6.4, we cover our feedback and observation. From the interaction design pattern of the motion gestures, we propose a novel interaction technique called TMR (Touch-Move-Release) for mobile AR.

3.6.1 Level of Agreement

As previously described in the working definitions of Wobbrock et al. [60], user-defined gesture sets are based on the largest set of identical gestures performed for a given task. In the process of gesture analysis, we found that similar gestures were proposed for the same task, and some were used across multiple tasks. We converted the elicited gestures into 504 points and combined similar gestures in each task. When combined the gesture into a group, we allocated the point accordingly and kept track of the score of each group of gestures. For example, in the *Slingshot* task, twenty participants proposed the same *Swipe-Down* surface gesture for it while only one participant proposed a *Tap* gesture. As a result, we classified two groups for the surface gesture in the *Slingshot* task, the former group with 20 points and the latter with 1 point.

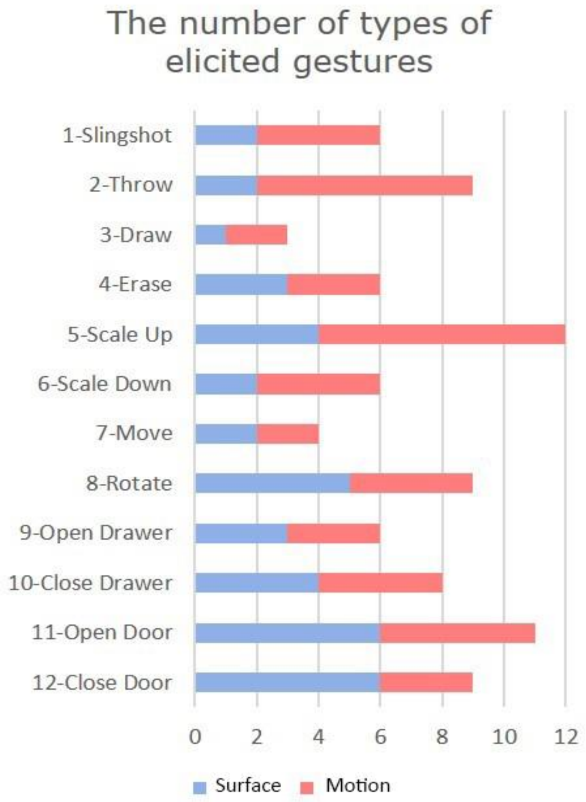


Figure 3: The number of types of gestures elicited for each task, surface gestures in blue and motion gesture in red.

We compared the level of consensus for each task. The agreement score was calculated using equation 1 for both sets of gestures based on [60].

$$A = \sum_{P_s} \left(\frac{|P_s|}{|P_t|}\right)^2 \qquad \text{- Equation 1}$$

P_t represents the total number of gestures elicited in the selected task, and P_s is the number of similar gestures categorised into the same group for that task. From our previous example, the surface gesture for the *Slingshot* task, which contained two groups of gestures with the scores of 20 and 1 points had an agreement score of 0.91, as shown in Equation 2. For the same task, the motion gestures had four groups of 14, 3, 2 and 2 points, and so the agreement score was found to be 0.48 as calculated in Equation 3.

$$A_S(\mathit{Slingshot}) = \left(\frac{|20|}{|21|}\right)^2 + \left(\frac{|1|}{|21|}\right)^2 = 0.91 \qquad \text{- Equation 2}$$

$$A_M(\mathit{Slingshot}) = \left(\frac{|14|}{|21|}\right)^2 + \left(\frac{|3|}{|21|}\right)^2 + \left(\frac{|2|}{|21|}\right)^2 + \left(\frac{|2|}{|21|}\right)^2 = 0.48$$ - Equation 3

For the *Throw* task, the calculation of the agreement scores for surface and motion gesture sets as follows:

$$A_S(\mathit{Throw}) = \left(\frac{|19|}{|21|}\right)^2 + \left(\frac{|2|}{|21|}\right)^2 = 0.83$$ - Equation 4

$$A_M(\mathit{Throw}) = \left(\frac{|8|}{|21|}\right)^2 + \left(\frac{|4|}{|21|}\right)^2 + \left(\frac{|2|}{|21|}\right)^2 + \left(\frac{|2|}{|21|}\right)^2 + \left(\frac{|2|}{|21|}\right)^2 + \left(\frac{|2|}{|21|}\right)^2 + \left(\frac{|1|}{|21|}\right)^2 = 0.22$$

- Equation 5

The results of agreement scores for each task were plotted and illustrated in Figure 4. The differences between the agreement scores of the two sets of gestures were notable for Task 1 - *Slingshot* (A_s=.91, A_m=.48) and Task 2 – *Throw* (A_s=.83, A_m=.22). While Task 9 - *Open Drawer*, Task 10 - *Close Drawer*, Task 11 - *Open Door*, and Task 12 - Close Door had low agreement scores for both sets of gestures. Based on the top consensus group of similar gestures in each task, we constructed two sets of user-defined gestures for the surface gesture set with 13 gestures and motion gesture set with 12 gestures, as shown in Figure 5.

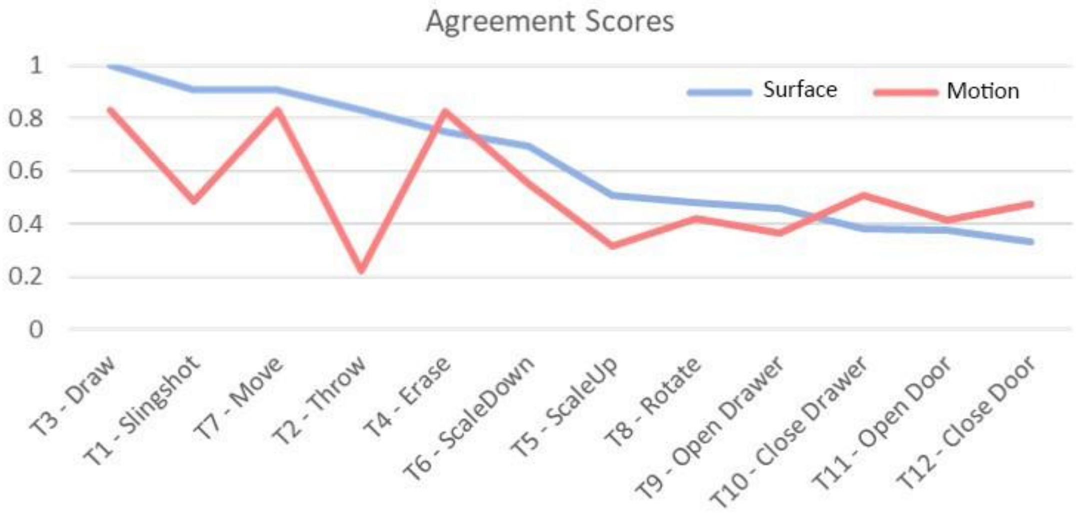


Figure 4: Agreement scores for each task in descending order, surface gestures (blue line) and motion gestures (red line).

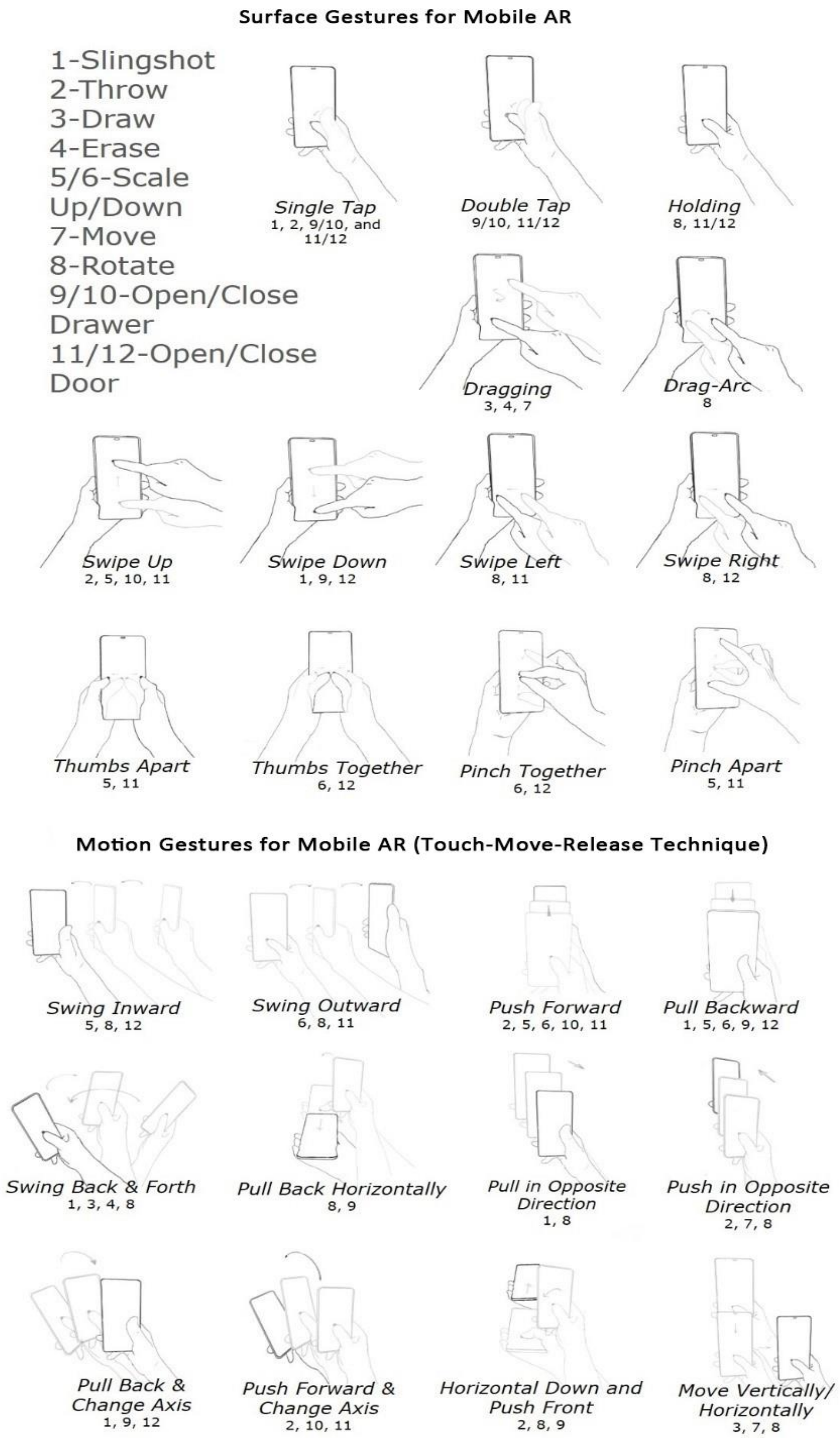


Figure 5: Two sets of user-defined gestures, surface gestures for mobile AR (top), motion gestures for mobile AR (bottom). Motion gestures demonstrate the concept of Touch-Move-Release (TMR) interaction technique.

3.6.2 User-defined Gesture Characteristics

From Figure 5, it is evident that each gesture can be used to perform multiple tasks. For example, in the surface gesture set, Task 9 - *Open Drawer*, 10 - *Close Drawer*, 11 *Open Door*, and 12 *Close Door* shared the *Double-Tap* gesture. Notably, *Swiping* and *Holding* gestures were also common occurrences across multiple tasks for the surface gesture set. For the motion gestures, we found that for some of the tasks, the agreement scores are lower than the surface gesture set (see Figure 4). This was our expectation as the motion gestures allow the handheld device movement in 3D space, which support greater possibilities in a larger design space.

Nevertheless, we observed some common characteristics and interaction pattern in the elicited motion gestures. Firstly, the trajectory of the gestures, i.e. device movement's direction, varied but was generally aligned with the desired movement of the manipulated virtual object. Secondly, participants could utilise the touch-sensitive screen to initiate and terminate their action. These observations led us to propose the Touch-Move-Release (TMR) technique, which involves three steps of action corresponding to the functions of initiating, performing and terminating an interaction.

3.6.3 Comparisons of Subjective Ratings

To validate our hypotheses, we analysed the three subjective rating scores between the surface and motion gestures in terms of *Goodness*, *Ease of Use*, and *Engagement*. We applied the Friedman test followed by a post-hoc pairwise comparison, Wilcoxon signed-rank tests with Bonferroni correction (with p-value adjusted), to compare the two sets of ratings. Figure 6, 7, 8 and 9 illustrate the plots for all the twelve tasks in terms of *Goodness*, *Ease of Use*, *Engagement*, and *Overall*. The *Overall* score was an average of all the three ratings. We used $\bar{x}_s$ to indicate the mean rating for the surface gestures and $\bar{x}_m$ for motion gestures, and the number in the bracket is the standard deviation.

Goodness Ratings: We found significant differences for Task 5 – *Scale Up* (V=123.5, p=.03, $\bar{x}_s$=6.1(1.2), $\bar{x}_m$=5.1(1.4)), Task 7 – *Move* (V=6, p=.03, $\bar{x}_s$=5.8(1.2), $\bar{x}_m$=6.4(0.6)), Task 9 – *Open Drawer* (V=91, p=.02, $\bar{x}_s$=6.1(1.0), $\bar{x}_m$=4.7(1.8)) and Task 10 – *Close Drawer* (V=75.5, p=.04, $\bar{x}_s$=6.1(0.9), $\bar{x}_m$=5.1(1.5)).

Ease of Use Ratings: Significant differences were found for Task 9 – *Open Drawer* (V=91, p=.001, $\bar{x}_s$=6.7(0.5), $\bar{x}_m$=5.3(1.6)), Task 10 – *Close Drawer* (V=101.5, p=.002, $\bar{x}_s$=6.6(0.5), $\bar{x}_m$=5.2(1.6)), Task 11 – *Open Door* (V=85, p=.04, $\bar{x}_s$=6.3(0.8), $\bar{x}_m$=5.7(1.1)), and Task 12 – *Close Door* (V=78, p=.02, $\bar{x}_s$=6.3(0.8), $\bar{x}_m$=5.4(1.3)).

Engagement Ratings: We found significant differences for Task 1 – *Slingshot* (V=9, p=.006, $\bar{x}_s$=5.1(1.3), $\bar{x}_m$=6.1(1.0)), Task 2 – *Throw* (V=39.5, p=.04, $\bar{x}_s$=5.0(1.2), $\bar{x}_m$=5.7(1.1)), Task 7 – *Move* (V=13.5, p=.005, $\bar{x}_s$=5.3(1.1), $\bar{x}_m$=6.4(0.7)), and Task 11 – *Open Door* (V=33.5, p=.04, $\bar{x}_s$=4.7(1.7), $\bar{x}_m$=5.8(1.2)).

Overall Score: The average yielded significant differences for Task 7 – *Move* (V=21.5, p=.02, $\bar{x}_s$=5.7(1.1), $\bar{x}_m$=6.4(0.7)), and Task 9 – *Open Drawer* (V=164.5, p=.03, $\bar{x}_s$=6.1(1.0), $\bar{x}_m$=5.0(1.7)).

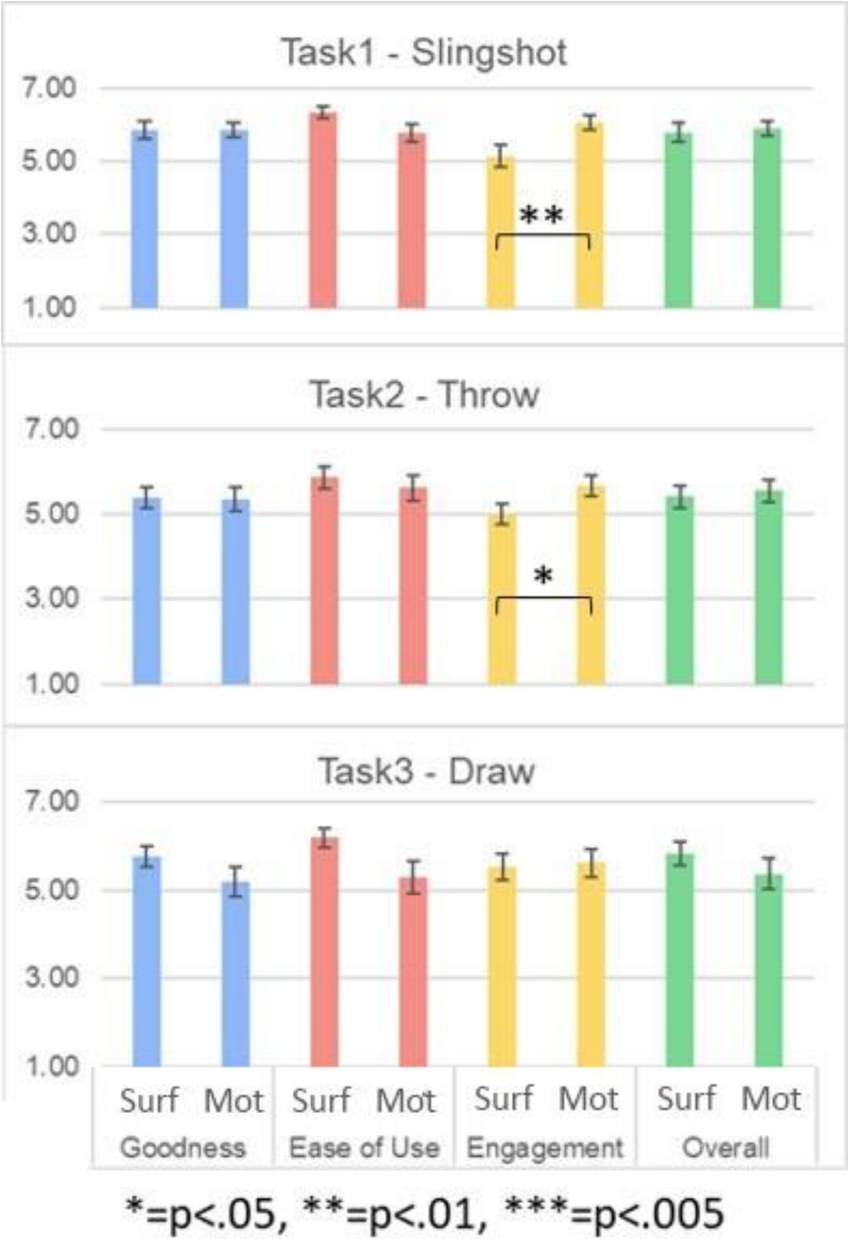


Figure 6: Subjective ratings in terms of *Goodness*, *Ease of Use*, *Engagement*, and *Overall* ratings for Task 1 to 3.

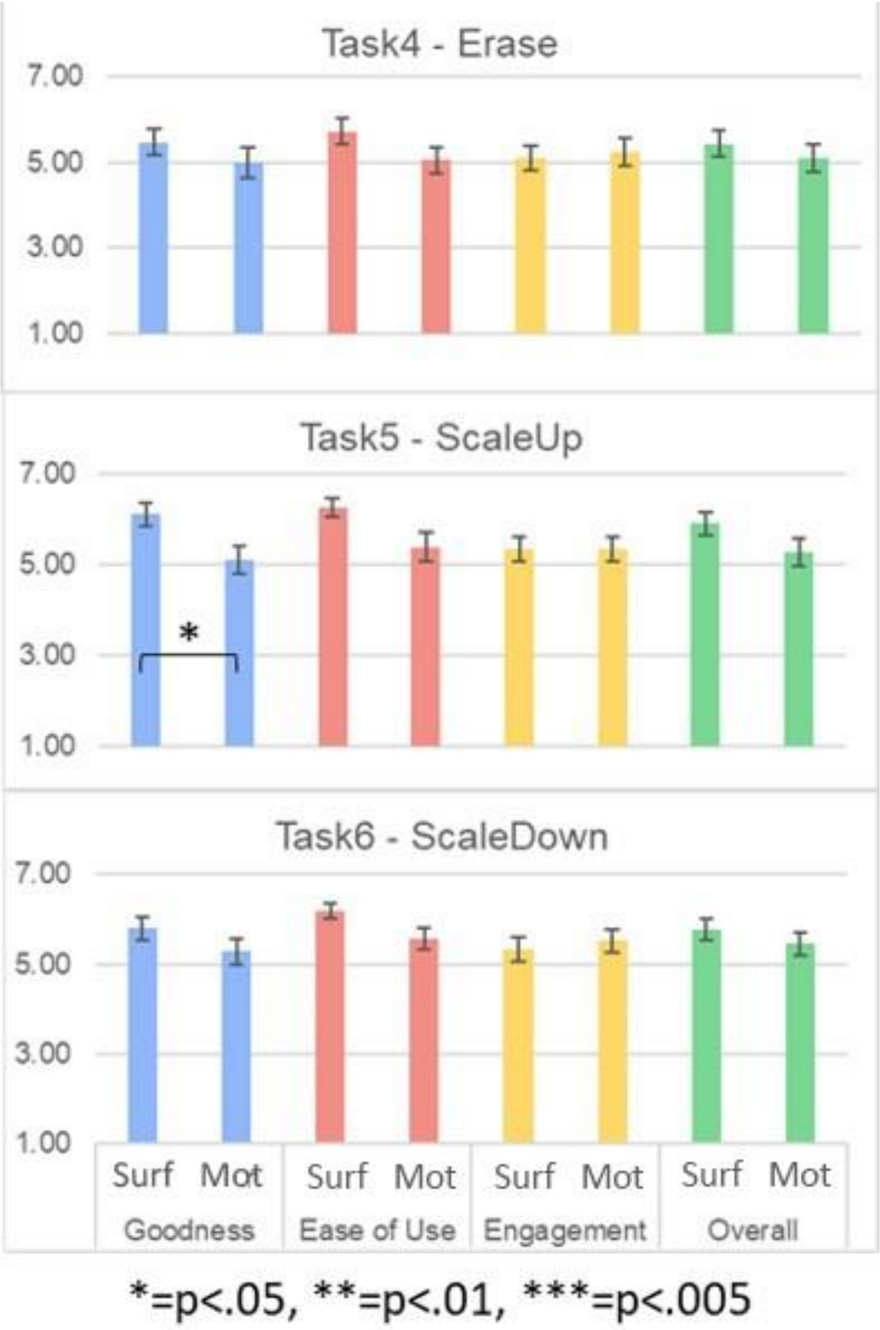


Figure 7: Subjective ratings in terms of *Goodness*, *Ease of Use*, *Engagement*, and *Overall* ratings for Task 4 to 6.

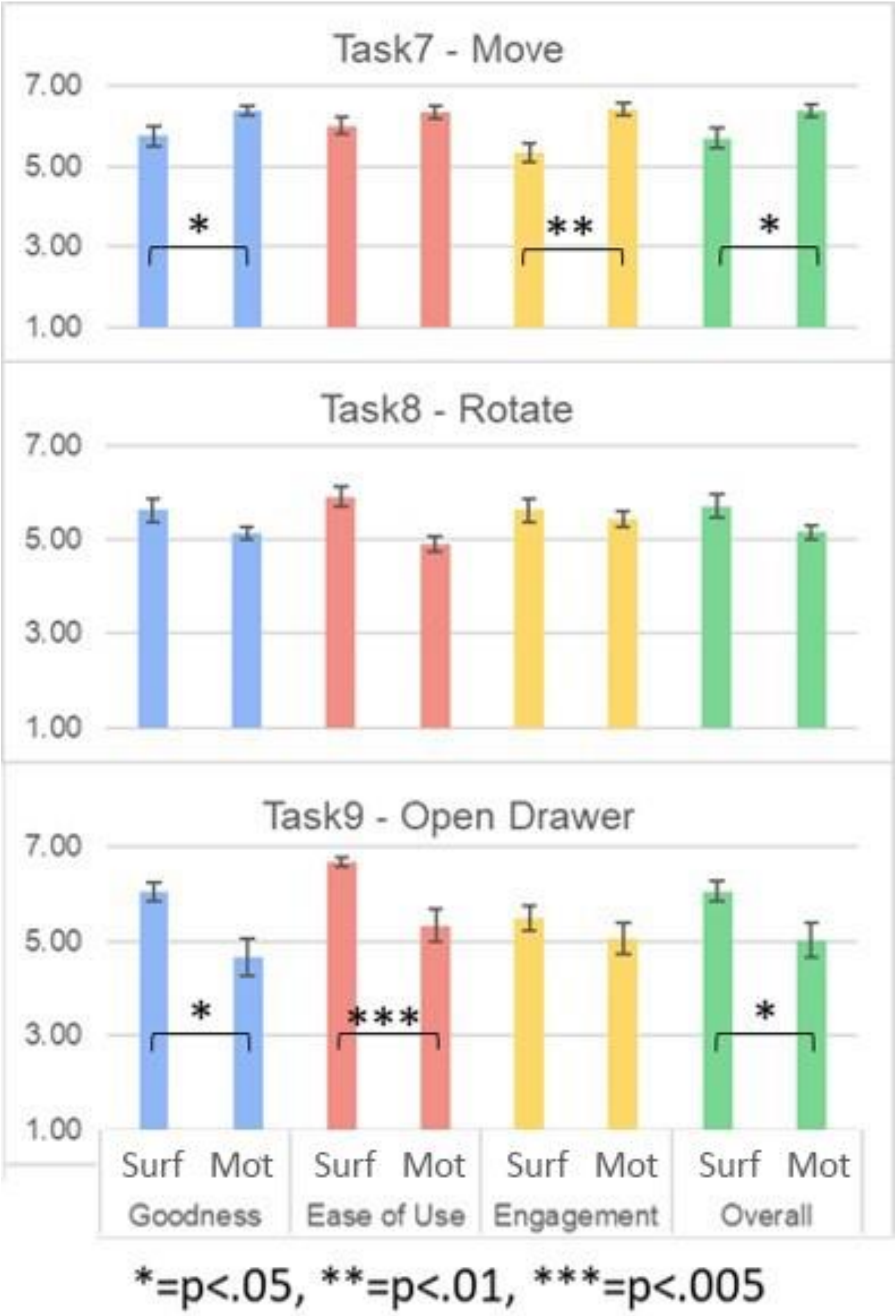


Figure 8: Subjective ratings in terms of *Goodness*, *Ease of Use*, *Engagement*, and *Overall* ratings for Task 7 to 9.

Figure 9: Subjective ratings in terms of *Goodness*, *Ease of Use*, *Engagement*, and *Overall* ratings for Task 10 to 12.

3.6.4 Feedback and Observation

During the elicitation process, we asked the participants to think aloud and explain their thoughts for their design decision and ratings of their gestures. At the end of the study, we also conducted a short unstructured interview for additional feedback. For example, we asked which gesture sets they would like to use in mobile AR applications. Thirteen participants chose motion gestures, while the remainders picked surface gestures. From the information collected, we identified design pattern and summarised into five themes, which found the gestures elicited to be reusable, multi-fingered or multi-trajectories, having trade-offs between ease of use and engagement, functionality-based, and finally context-based.

Reusable – Many gestures elicited could be used for multiple tasks. These gestures are more common for the surface gestures set, for example, *Tap* gesture could be applied for Task 1 – S*lingshot* (1 vote), Task 2 – *Throwing* (2 votes), Task 4 – *Erase* (2 votes), Task 5 – *Scale-Up* (1 vote), Task 9 – *Open Drawer* (5 votes), Task 10 – *Close Drawer* (10 votes), Task 11 – *Open Door* (12 votes), Task 12 – *Close Door* (11 votes). The other gestures that were considered reusable included *Double-Tap*, *Long Press*, *Swipe Up*, *Swipe Down*, *Swipe Left*, and *Swipe Right*. In comparison, the motion gestures were not as common; some of the reusable ones were *Tap-Phone*, *Swing-Release* and *Tap-Pull Back-Release*. The same gesture should be applicable for multiple tasks as it might help lower user's mental effort having to learn and recall a fewer number of gestures.

Multi-fingered or Multi-trajectories – We observed that some participants paid attention to how many fingers they were using for the surface gesture. Some participants were also trying to minimise the number of fingers used. For instance, in task 6 – *Scale Down*, seven participants initially designed the gesture with five fingers, *Pinch Together*, on the screen but then changed to the conventional two-finger pinch. A possible reason might be due to the finger's occlusion of the screen, as stated by *Participant 4* as follows:

> *"I feel that in surface interaction, the extra fingers will obstruct my screen view."* – P_4

Two participants also asked if the UI interface could be provided, so they could directly tap the on-screen button to scale instead. Below was a comment made by *Participant 14* on providing UI to assist with the interaction.

> *"Can I imagine a slide bar button in the scene? When I slide the button, the virtual object will scale down automatically."* – P_{14}

For the motion gestures, we observed that all the participants performed the gestures using a single hand, their right hand, and only use their thumb for input on the touch screen. The participants mainly focused on working out the appropriate movement trajectories of the handheld device for the motion gesture design.

Trade-Offs Between Ease of Use and Engagement – When the participants were asked to rate their gestures in terms of *Ease of Use* and *Engagement*, they would consider the duration of the interaction. Some participants felt that for prolonged usage of surface gestures might be less fatiguing and more efficient than motion gestures for certain tasks. *Participant 15* mentioned that moving the device around could be quite tiring:

> *"Holding the phone for a long time makes my palms sweat, and more physical movements will exacerbate the situation."* – P_{15}

Nevertheless, some participants felt that surface gestures could be quite boring to use for a long duration, and motion gestures could deliver a better experience for some tasks in 3D space as *Participant 17* pointed out that:

> *"I have played Pokémon Go before… I like its story more than swiping up the screen to capture the Pokémon."* – P_{17}

Functionality-based – Some participants gave different subjective ratings for dichotomous tasks with similar gestures. For example, *Participant 5* rated Task 9 – *Open Drawer* for *Double Tap* surface gesture, **7 / 7 / 7** (Goodness/Ease of Use/Engagement), while rated *Tap-Pull-Release* motion gesture, **1 / 2 / 3**. On the contrary, *P5* rated Task 10 – *Close Drawer, Swipe Up* for surface gesture, **4 / 6 / 5** and *Tap-Down-Forward-Release* motion gesture, **6 / 6 / 6**. *P5* explained that when opening the drawer using the surface gesture, the content inside the drawer could be seen immediately, however, with the motion gesture, the camera might lose the view of the drawer during the manipulation. Nevertheless, when the drawer needed to be closed, the content inside the drawer might not be important. Therefore, *P5* found motion gestures more engaging to use.

Context-based – Some participants felt that their ratings would depend on the context of the application, for example, if it is in gaming context then the motion gestures might be rated highly for *Goodness* but might be rated much lower in non-entertainment applications. Therefore, the application context should be taken into account when choosing to use the surface or motion gesture sets

> *"While playing a mobile AR game, I feel motion gestures have an irreplaceable charm, like the Joy-Con of Nintendo Switch, it will bring more realistic user experience. But for some applications scene that requires accuracy, surface gestures are more appropriate."* – P_1

3.7 Summary

In this chapter, we have found our answers to the first two research questions. The first question led us to our first goal to find gesture-based interaction that could utilise 3D space that mobile AR applications support for a handheld device. Through an elicitation study, we have elicited both the surface gesture set and the motion gesture set. From the elicited motion gestures, we have discovered an interaction pattern and proposed a technique where the participants could use the device’s touchscreen to initiate and terminate interaction and device’s AR tracking for device’s movement to engage the user with more physical activity in 3D space for better AR experience. We called this technique, TMR (Touch-Move-Release). For the second research question, we were interested in how the participants would perceive motion gestures comparing to the conventional surface gestures. Therefore, during the elicitation process, we also collected subjective ratings in terms of *Goodness*, *Ease of Use*, and *Engagement*. There were a number of interesting findings. We will further discuss our results in Section 5.1.

Chapter 4

Validation Study

In this chapter, we present our findings from investigating the final research question, which asks how the proposed motion gestures, TMR (Touch-Move-Release), would fair against the conventional surface gestures in an actual mobile AR application. We chose to implement a game modelled after Pokémon Go for this comparison. Pokémon Go is, arguably, the most popular mobile AR game at the time of this writing, released by Niantic in 2016. The main goal of the game is for the players to collect various Pokémon, virtual creatures with different abilities, which are procedurally generated and scattered using geospatial information throughout the real world [39].

In the actual game, to catch a Pokémon, the player has to aim and perform an onscreen *Swipe Up* gesture to throw a virtual Pokéball to capture a Pokémon. The ball's velocity (speed and direction) is controlled by the player's finger sliding from the bottom toward the top of the screen. For the comparison, we implemented the gestures elicited for Task 2 – *Throw* from the elicitation study, where a *Swipe Up* gesture was one of the surface gestures elicited. We chose this gesture to be our baseline condition in this evaluation. For the motion gesture, we implemented the *Push Forward & Change Axis* gesture. We will give an overview of the development of our system in Section 4.1, study design in Section 4.2, our participants in Section 4.3, the procedures in Section 4.4, our hypotheses in Section 4.5, the results of the study in Section 4.6, and finally summarise this chapter in Section 4.7.

4.1 System Prototyping

Our Pokémon GO clone prototype could be played indoor or outdoor similar to the original game, utilising the geospatial information of the current player's location. The goal of our game was for the player to capture and complete a collection of the Pokémon-like creatures. As our focus was not on the gameplay but to compare the two gesture interaction techniques, we did not create any fixed point in the real world for the creature to spawn. On the contrary, we created a menu for the player to spawn the creatures around their current location. Only when the player was within the proximity threshold (i.e. 2 meters), the creatures became visible on the handheld device screen, allowing them to throw a Pokéball-like ball to capture them. Once captured, those creatures would be indexed in the Pokédex-like encyclopedia for the player to keeping track of their creatures. We used the Unity Engine [54] with the Vuforia SDK [45] for the development of the Pokémon GO clone game. Screenshots of our game and further descriptions are given in Figure 10.

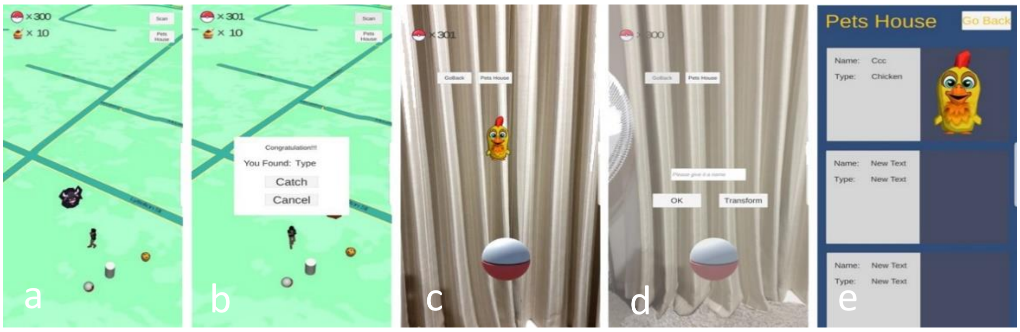


Figure 10: Screenshots of the Pokémon Go clone game: a) World map showing the player's current geolocation in the real-world and the locations of the creatures, b) the player is prompted to proceed for a catch once within the vicinity of the Pokémon, c) the creature appears in the AR view, registered in the real-world 2 meters in front of the device camera, d) the creature is captured, e) the encyclopedia shows the information of the captured creature.

4.1.1 Unity Engine

Unity Engine [54] is a popular game development engine, developed by Unity Technologies. It supports cross-platform game development, enabling the developers to build 2D and 3D applications through scene authoring and the concept of a tree hierarchical data structure with GameObjects as the nodes in the scene tree. The GameObjects contain various components, e.g. attributes, scripts etc. For our prototype development, we used the Unity Engine version 2017.3.1f1 and the C# programming language for scripting the GameObject's behaviours.

4.1.2 Vuforia Software Development Kit

Vuforia SDK [45] is a software development kit that supports the creation of mobile AR experiences. It was initially developed by Qualcomm [46] and later acquired by PTC Software [44] in 2015. Vuforia is able to determine the position and orientation of the device in 3D space and identifying the ground plane in the real-world using the mobile device's camera. We used the Vuforia SDK (version 8.0) for Unity to develop our mobile AR Pokémon GO clone game. With Vuforia SDK, our system was able to place virtual creatures in the real world, and the device's position and orientation were tracked relative to the physical environment, providing the device's spatial localisation ability.

4.1.3 Location-Based Service

Beyond the localisation of the device relative to the immediate surrounding, we also integrated the location-based services (LBS) [50] in our game to replicate the experience of the original Pokémon Go. LBS utilises the player's geographic location with the general notion of services, e.g. dialling an emergency number or a navigation system. We illustrate the context of LBS in Figure 11. During the gameplay, the mobile device obtains the real-time location information of the player through the network provider and global positioning system (GPS). Combining LBS and Vuforia SDK, we were able to accurately determine the device's location in 3D space outdoor as well as indoor to the precision of the underlying technology.

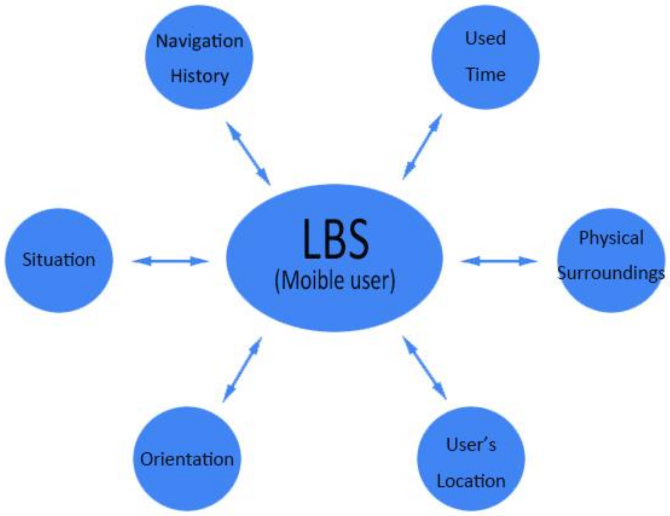


Figure 11: Context of location-based service (LBS).

4.1.4 Gestures Implementation

For both gestures, the ball throwing action was implemented based on Unity's physics simulation. The ball (Pokéball) was implemented as a spherical rigid body that was influenced by the gravity of the physics simulation and followed a projectile motion.

Baseline Surface Gesture (*Swipe Up*) – Our system registered and recorded the initial position when the player first touched the ball on the screen. As the player slid the finger upward and lifted off the screen, the last position and the duration of the touch were recorded. The difference between the final and the initial touch position yielded a movement vector of the finger. We were dividing the movement vector by the time duration, which gave us the velocity (magnitude and direction) of the finger. A force vector was created based on a chosen constant times the velocity vector at the ball's position. As the force was only exerted after the Swipe Up gesture was performed, there was a small delay between the player's action and the ball being thrown. However, the gesture was typically fast, and the lag was not noticeable to the player. Furthermore, the finger also occluded the screen, and so it was too difficult to notice the delay.

TMR Motion Gesture (*Push Forward & Changing Axis*) – For TMR, we took into account of two factors, the first factor was the time when the thumb was lifted off the screen (Release), and the second was the angular velocity of the device's motion being pushed forward. To track the device's movement, we obtained the data from the inertial measurement unit (IMU), which combined an accelerometer and a gyroscope to obtain the orientation and linear acceleration of the mobile device. We set the update rate of the IMU data at 0.1 seconds. The initial device's pose was typically almost vertical to the ground. The player first registered the touch on the screen with a thumb to simulate holding onto the ball. Then the player swung the phone forward with some velocity and then released the thumb to throw the ball. When the player first touched the screen, the system started monitoring the angular velocity of the device, looking for any change in angular velocity (tilt forward with speed) and if that change exceeded the threshold. The linear force was calculated based on the linear velocity of how far and how fast the device moved forward during the movement. The calculation was similar to the surface gesture but with the device's movement instead of the finger's movement.

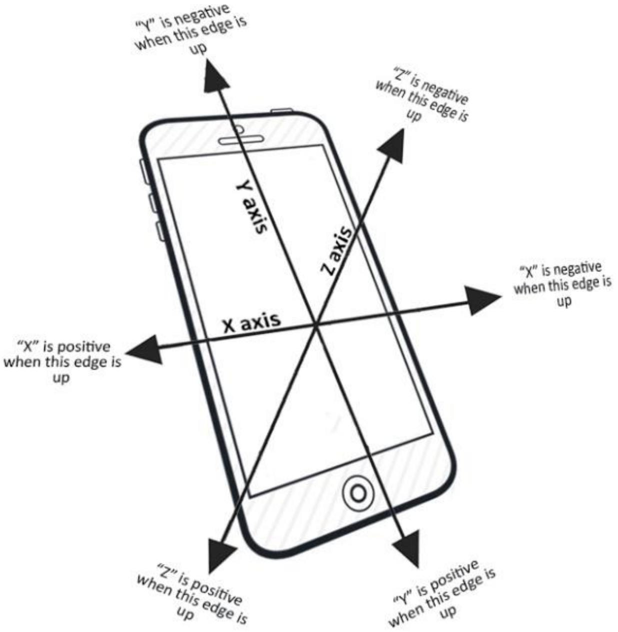


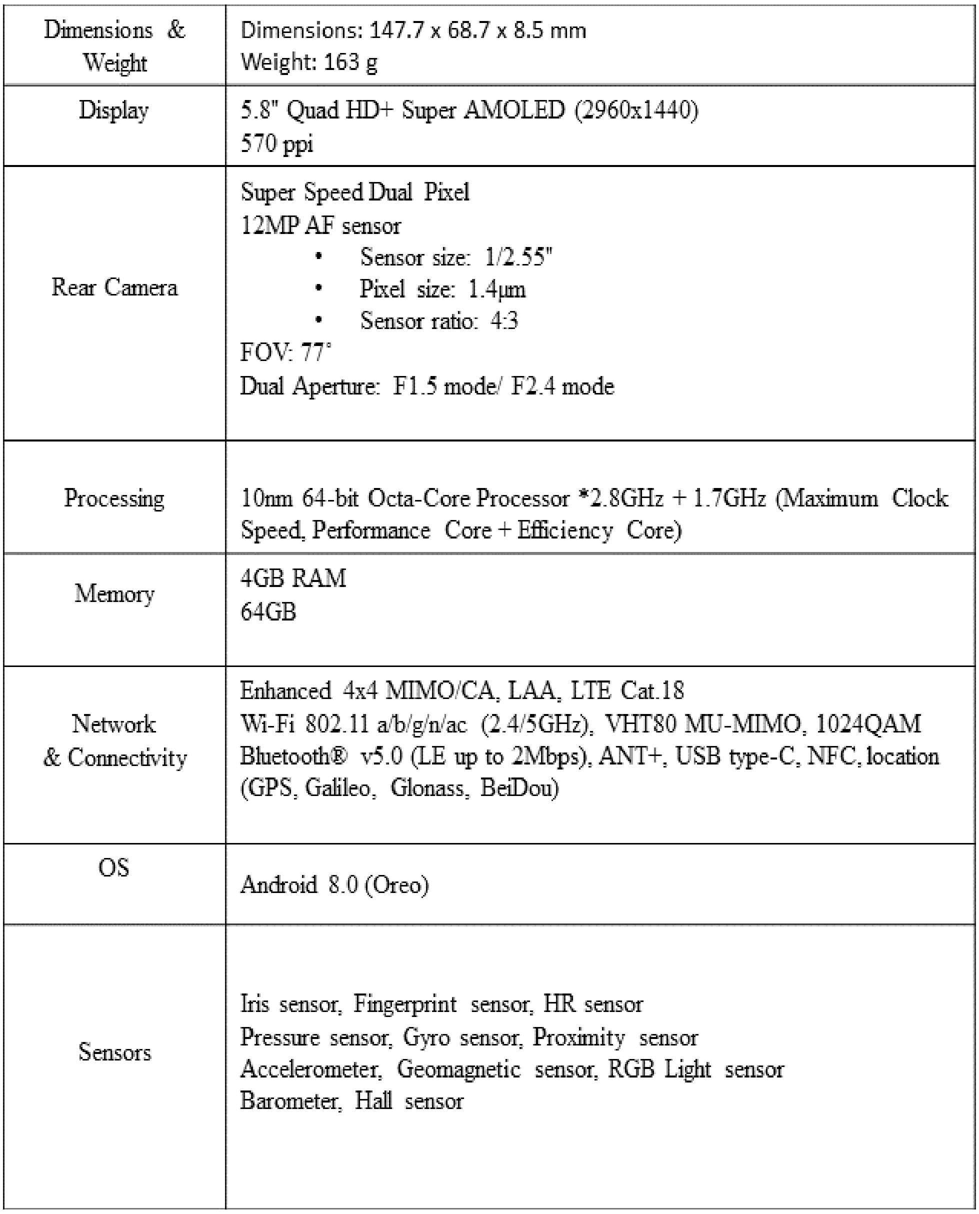
Figure 12: Accelerometer's linear acceleration along each axis.

4.1.5 Device Hardware

For this study, we chose the Samsung Galaxy S9 mobile device. Samsung Galaxy S9 is an Android-based smartphone which comes with the hardware specification shown in Table 2. It satisfied the requirement for running our

Pokémon GO clone game, which required high mobile processing power, built-in sensors, a camera, and supported Android 9.0 Pie.

Table 2: Samsung S9 Specifications



Dimensions & Weight	Dimensions: 147.7 x 68.7 x 8.5 mm Weight: 163 g
Display	5.8" Quad HD+ Super AMOLED (2960x1440) 570 ppi
Rear Camera	Super Speed Dual Pixel 12MP AF sensor • Sensor size: 1/2.55" • Pixel size: 1.4µm • Sensor ratio: 4:3 FOV: 77° Dual Aperture: F1.5 mode/ F2.4 mode
Processing	10nm 64-bit Octa-Core Processor *2.8GHz + 1.7GHz (Maximum Clock Speed, Performance Core + Efficiency Core)
Memory	4GB RAM 64GB
Network & Connectivity	Enhanced 4x4 MIMO/CA, LAA, LTE Cat.18 Wi-Fi 802.11 a/b/g/n/ac (2.4/5GHz), VHT80 MU-MIMO, 1024QAM Bluetooth® v5.0 (LE up to 2Mbps), ANT+, USB type-C, NFC, location (GPS, Galileo, Glonass, BeiDou)
OS	Android 8.0 (Oreo)
Sensors	Iris sensor, Fingerprint sensor, HR sensor Pressure sensor, Gyro sensor, Proximity sensor Accelerometer, Geomagnetic sensor, RGB Light sensor Barometer, Hall sensor

4.2 Study Design

In this study, we conducted a 2×3 within-subjects factorial design study where the two independent variables were the two interaction techniques (surface and motion gestures) and the three target creature sizes (*Small*, *Medium*, and *Large*). The combination yielded six conditions, as shown in Table 3. The two interaction techniques were counterbalanced, and the three sizes were then counterbalanced for each technique. The dependent variables were the accuracy of the throw by counting the number of balls, i.e. Pokéballs, to capture the creature used before successfully hitting the target. Between the two interaction techniques, we also collected and compared their system usability scale (SUS) [5], game experience questionnaire (GEQ) [29], and subjective ratings on *Goodness*, *Ease of Use*, and *Engagement* as in the elicitation study.

The task was for the participants to play our game to capture the creatures. To complete the game, the participant must capture two target creatures of each size (*Small*, *Medium*, and *Large*), as shown in Figure.13. The target creature would be automatically spawned when the player is within a two-meter radius of the indicated location of the creature. The target creature would be placed 2 meters in front of the player's device location. We recorded how many balls the participants used in each condition. We asked our participants to answer the three sets of questionnaires (SUS, GEQ, and subjective ratings) after completing each interaction technique. A short unstructured interview was conducted at the end of the study.

Table. 3: Conditions of Gesture Validation

Target Size / **Interaction Techniques**	**Small (height=30 cm)**	**Medium (height=60 cm)**	**Large (height=120 cm)**
Surface Gesture (Baseline)	*Swipe Up –* 30 cm target	*Swipe Up –* 60 cm target	*Swipe Up –* 120 cm target
Motion Gesture (TMR)	*Push Forward –* 30 cm target	*Push Forward –* 60 cm target	*Push Forward –* 120 cm target

4.3 Participants

This study was filed under the same ethics application reference number of HEC 2019/94/LR approved by the Human Ethics Committee, the University of Canterbury on the 14th of November 2019. The participants had to sign a consent form, which contained the experiment information. The participants were told that they could discontinue the experiments at any time without penalty regardless of the fact that there was not any serious health and safety concern. The participation was voluntary, and the participants were told that this was an opportunity for them to try a mobile AR experience. We recruited 10 participants (six females) with an age range of 18 to 40 years old ($\bar{x} = 26.4$, SD = 5.95). All participants owned a smartphone and used the touchscreen every day. They are highly familiar with conventional surface gestures. Five participants had no previous experience with mobile AR applications, and the other remaining participants had limited experience with the original Pokémon GO game.

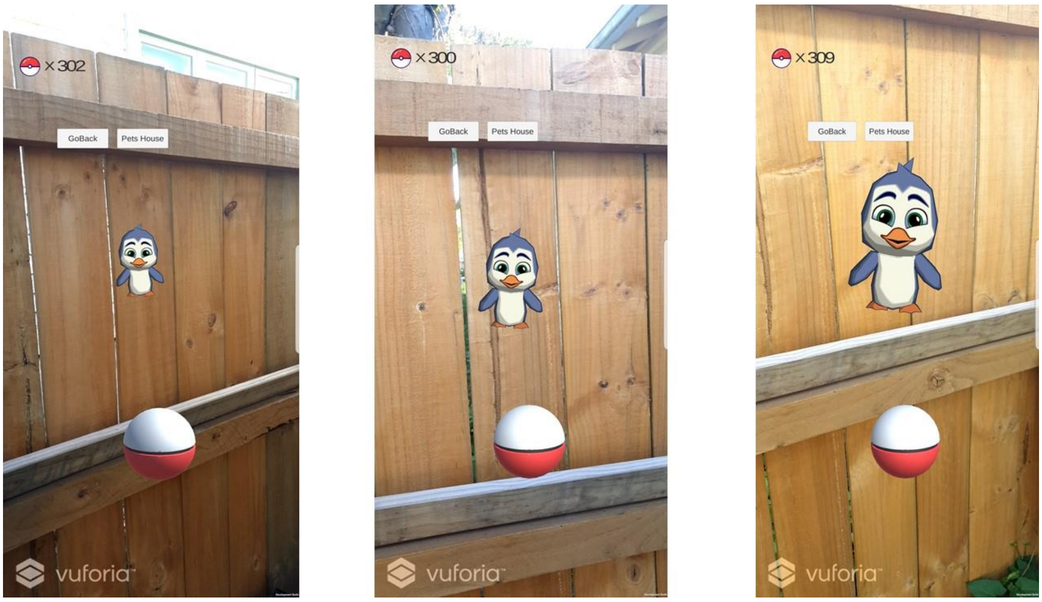


Figure. 13: Three size of Pokémon-like creatures.

4.4 Procedures

The procedures of this study were as follows:

1. We informed the participants of the study and asked them to read the information sheet as well as signing the consent form as shown in Appendix B.1 and A.2
2. The participants answered the pre-experimental questionnaire (see Appendix A.3) to collect demographics information and their previous experience with mobile AR applications.
3. For training, the participants had at least two minutes or longer to familiarize themselves with our Pokémon GO clone game and UI. During this time, we encouraged them to ask any question around the operation of this application.
4. The participants were asked to capture six creatures, three of each size (*Small*, *Medium*, and *Large*) using one of the interaction techniques based on the counterbalancing of the conditions. The number of balls being thrown until the creature was captured, was recorded.
5. When completed one of the interaction techniques, the participants were asked to complete the subjective rating questionnaire (see Appendix B.3), the in-game GEQ (see Appendix B.2), and SUS (see Appendix B.4).
6. Step 4 and 5 were repeated for the other interaction technique.

7. As the participants completed all the tasks, they were asked to complete the post-experiment questionnaire (see Appendix B.5) to choose their preferred gesture and comment on the overall experience of the game.

4.5 Hypotheses

In this study, we compared the two interaction techniques, a baseline surface gesture and a TMR motion gesture on a throwing task in terms of accuracy with varying target sizes, subjective ratings, usability, and the overall experience. Our hypotheses were **there would be significant differences in accuracy *(H1)*, subjective ratings *(H2)*, in-game experience (*H3*), and system usability *(H4)* between the two interaction techniques**.

4.6 Results

We present the results of the comparison between the two interaction techniques and their accuracy for various target creature sizes in Subsection 4.6.1, the subjective ratings on goodness, ease of use, and engagement in Subsection 4.6.2, the scores of the system usability scale in Subsection 4.6.3, the outcomes of the game experience questionnaire in Subsection 4.6.4, and the preferences of the participants between the two gestures for the chosen task in Subsection 4.6.5.

4.6.1 Accuracy with Various Target Sizes

For the *Small* targets, the participants used a total of 51 balls ($\bar{x}$ = 5.1, SD = 2.3), and 85 balls ($\bar{x}$ = 8.5, SD = 5.4) for the baseline *Surface* and *TMR* motion gesture, respectively; 23 balls ($\bar{x}$ = 2.3, SD = 0.61) and 57 balls ($\bar{x}$ = 5.7, SD = 2.3) for the *Medium* targets; 21 balls ($\bar{x}$ = 2.1, SD = 0.3) and 37 balls ($\bar{x}$ = 3.7, SD = 2.1) for the *Large* targets. Figure 14 and Figure 15 illustrate the number of balls used for each target size by all the participants for each interaction technique.

We ran two-way ANOVA to examine the effect of interaction techniques and the target sizes on the accuracy of the throwing task. Figure 16 illustrates the plot of the number of balls used for each condition. We could not find any significant interaction effect between the two variables, *interaction techniques × target sizes*, on the accuracy of the throw. We found that *Swipe Up* gesture was significantly more accurate than *Push Forward & Changing Axis* gesture ($F_{1,59}$=14.7, p=.0003). We also found that there were significant differences between target sizes ($F_{2,59}$=10.1, p=.0002).

For the post-hoc pairwise comparisons, we used Wilcoxon signed-rank (WSR) tests with Bonferroni correction (p-value adjusted). We found that there were significant differences for surface gesture between *Small-Medium* targets (W=85, p=.006), and *Small-Large* targets (W=78, p=.002), and for motion gesture between Small-Large targets (W=116.5, p=.01) and Medium-Large (W=95.5, p=.04).

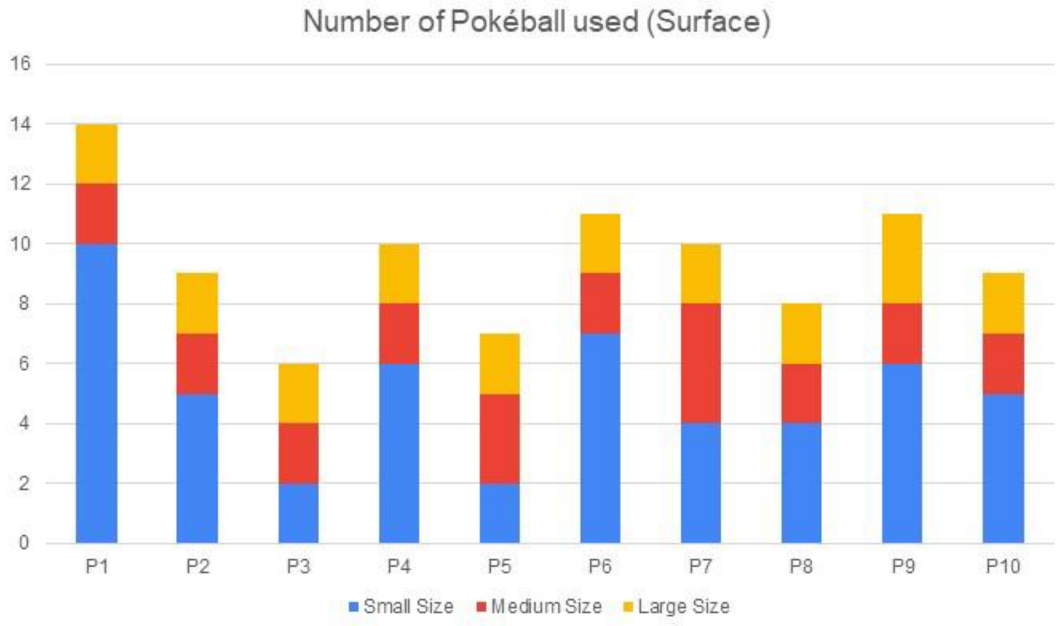


Figure 14: Number of balls used in the *Surface* gesture conditions.

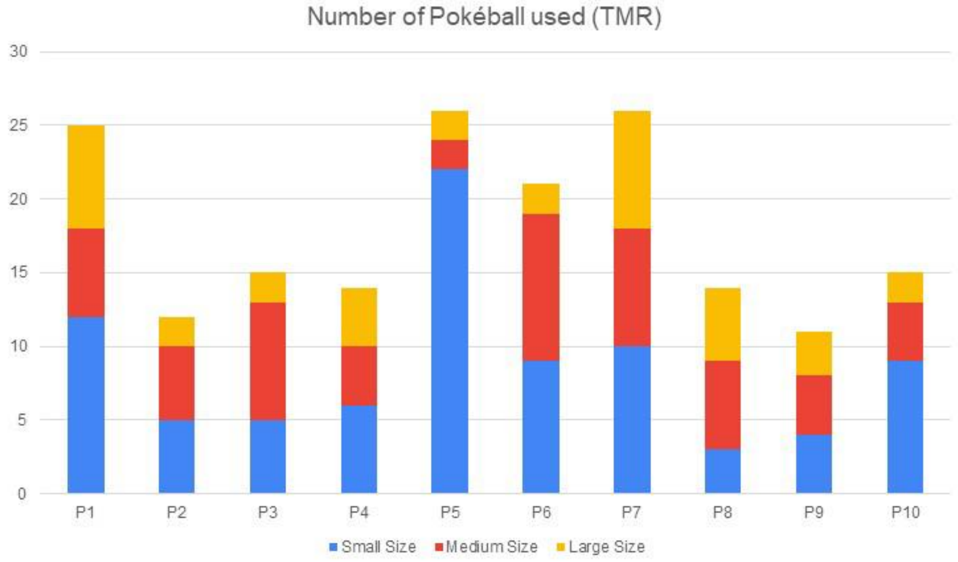


Figure 15: Number of balls used in the *TMR* gesture conditions.

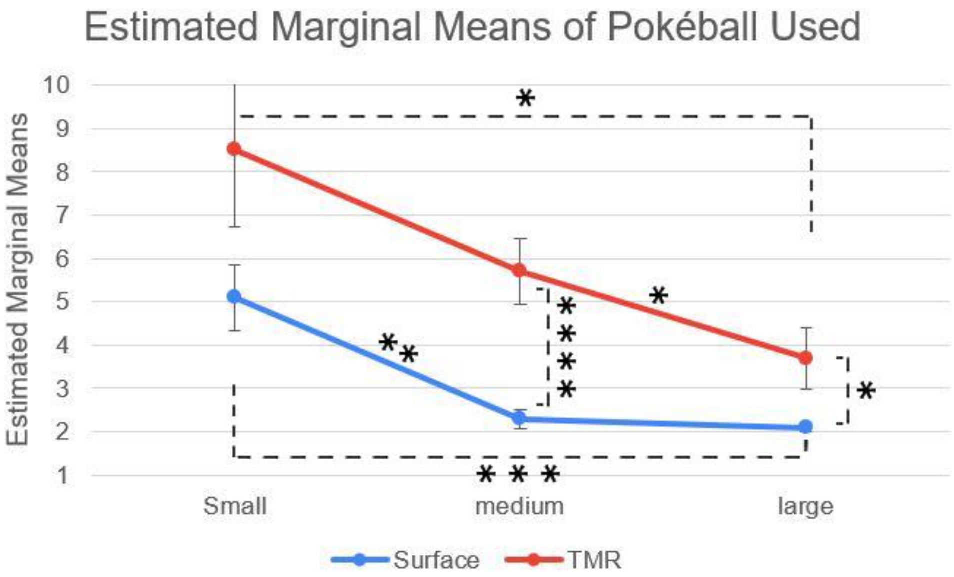
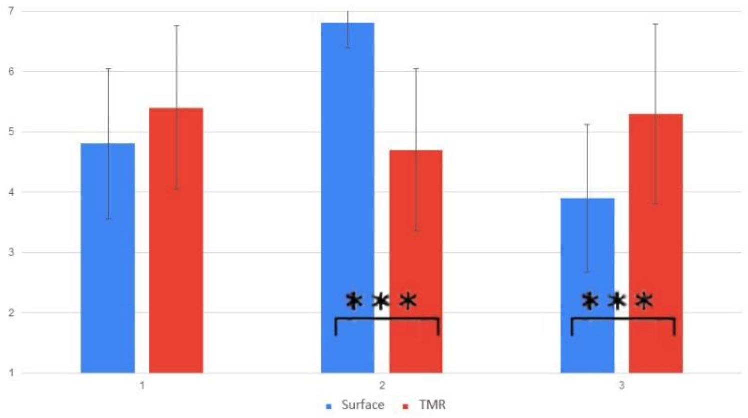


Figure 16: Number of balls used in each condition (*=p<.05, **=p <.01, ***=p <.005, ****=p <.001).

4.6.2 Subjective Ratings: Goodness, Ease of Use, and Engagement

Similar to the previous elicitation study, we asked the participant to rate the interaction based on three aspects in terms of *Goodness*, *Ease of Use*, and *Engagement*. For *Goodness*, the *Surface* gesture was rated x̄ = 4.8 (SD=1.3) and TMR, x̄ = 5.4 (SD=1.4). For Ease of Use, *Surface* gesture was rated x̄ = 6.8 (SD=0.4) and *TMR* gesture, x̄ = 4.7 (SD=1.3). Lastly, *Engagement*, Surface gesture scored x̄ = 3.9 (SD=1.2) and *TMR* gesture, x̄ = 5.3 (SD= 1.5). Figure 17 illustrates the average ratings for the two interaction techniques. We applied WSR tests and found significant differences for *Ease of Use* (W=36, p=.01) in favour of the *Surface* gesture, and for *Engagement* (W=0, p=.03) in favour of *TMR*.



*=p<.05, **=p<.01, ***=p<.005

Figure 17: Subjective ratings between the two interaction techniques.

4.6.3 Game Experience

We collected the participants feedback on the game experience scale for both in-game and post-game experiences. For the in-game experience scale, there were seven subscales on *Competence*, *Sensory and Imaginative Immersion*, *Flow*, *Tension*, *Challenge*, *Negative Affect*, *Positive Affect* from 14 questions (see Appendix B.2). With WSR tests, we found that TMR was rated significantly higher for ***Competence*** (*TMR* – x̄=3.1 (SD=0.4), ***Surface*** – x̄=2.3 (SD=1.2), W=5, p=.04), ***Sensory and Imaginative Immersion*** (*TMR* – x̄=2.6 (SD=0.6), *Surface* – x̄=2.4 (SD=1.2), W=2, p=.04), ***Flow*** (*TMR* – x̄ = 2.1 (SD=0.9), *Surface* – x̄ = 1.6 (SD = 1.1), W=5, p=.04), and ***Challenge*** (*TMR* – x̄=2.9, SD=0.9, *Surface* – x̄=0.9, SD=0.7, W=0, p=.006). There was no significant difference found for *Tension*, *Negative Affect*, and *Positive Affect*. Figure 18 illustrates the in-game experience plot. One of the participants gave an opinion on the reason for the difference for the *Challenge* rating as follows:

> *"TMR interaction is quite challenging for me, especially when the size of the Pokémon is extremely small."* – P5
>
> *"TMR interaction gesture can bring a sense of victory for me. I am proud of myself when I completed the task of capturing small size Pokémon." – P5*

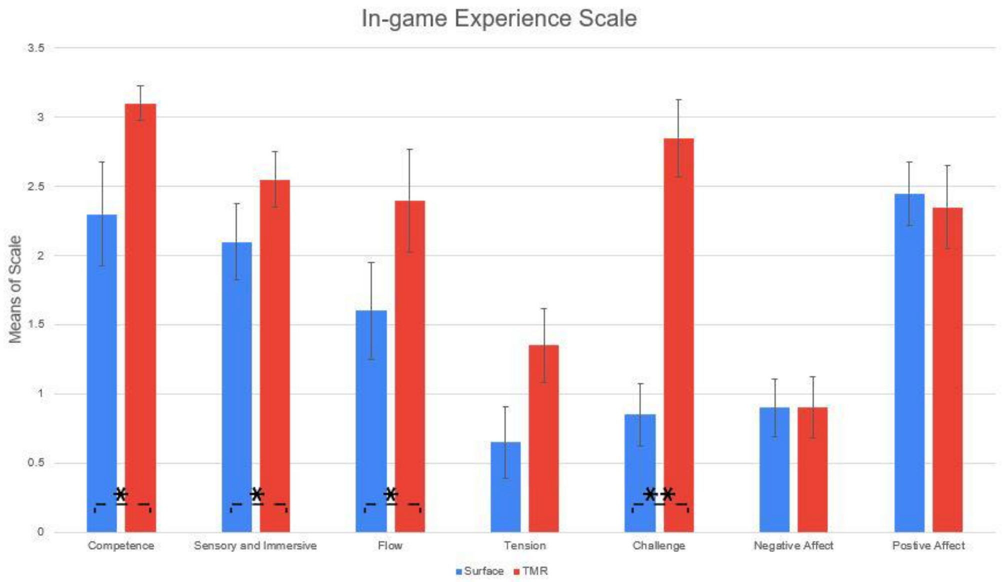


Figure 18: In-game experience scale. (*=p <.05, **=p <.01)

4.6.4 System Usability Scale

In terms of system usability scale (SUS), the surface gesture received the scores between 62.5 to 95 with the average of x̄=77 (SD=9.9), and TMR's scores ranged from 37.5 to 87.5 with x̄=68 (SD=15.9). Figure 19 shows the plot of all the SUS scores. A WSR test yielded a significant difference (W = 28, p-value = 0.02)

> *“Both interaction methods can complete the task. I feel good for the whole cycle of the demo. I think I need more levels to feel challenged.”*
>
> *– P8*
>
> *“TMR interaction lets my vision off the phone screen. If I try to capture the screen with my eyes, it makes me dizzy. Honestly, I prefer simple surface interaction.” – P3*

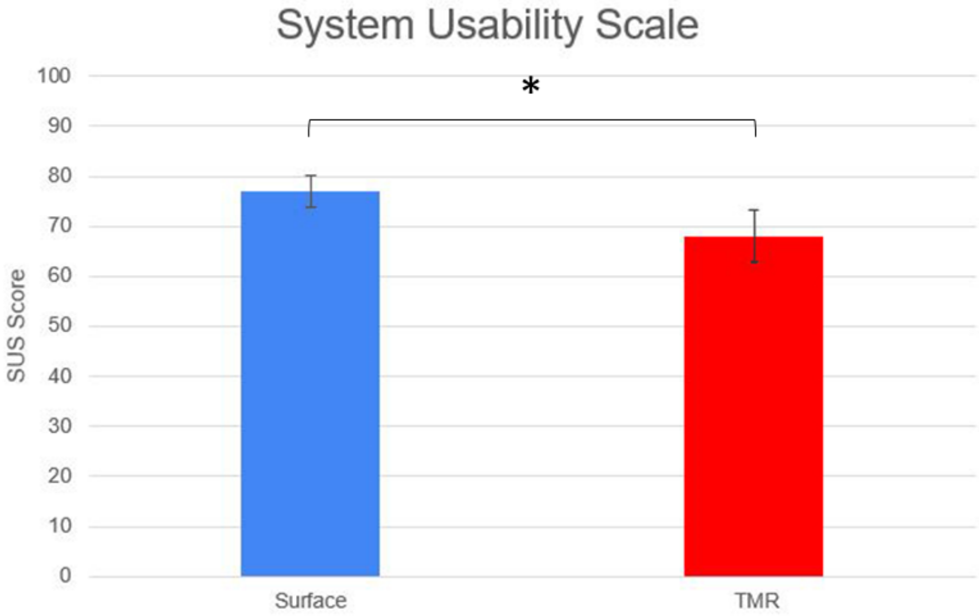


Figure 19: SUS scores for the two interaction techniques.

4.6.5 Preferences

From the post-experiment questionnaire, we asked our participants, which interaction technique that they preferred, half of them preferred the baseline *Surface* gesture, while the other half chose the *TMR* motion gesture, as shown in Figure 20. Those who preferred *Surface* gesture found it to be easier to use, while those who liked *TMR* gesture found it more challenging and felt more engaged with the game.

> *“I have been using surface interaction for a long time, and it is a habit for me. TMR interaction has unnecessary physical consumption for me.” – P7*
>
> *“The TMR interaction method makes me feel a challenge when using every Pokeball, which greatly increases the fun of the game. I hope that more complex and diverse TMR gestures can be used in the future.” – P2*

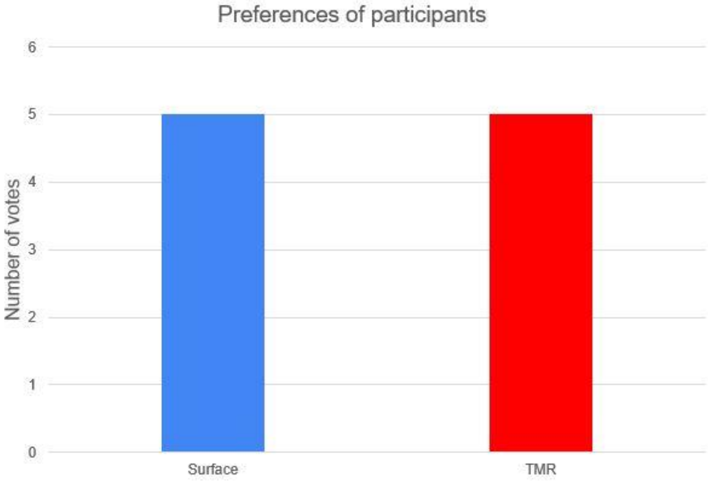


Figure 20: User preference of the interaction techniques.

4.7 Summary

In this chapter, we presented our findings from a validation study to compare the two types of interaction techniques, a baseline *Surface* gesture and a *TMR* motion gesture, which we elicited in the previous study. We developed and shared an overview of the mobile AR game similar to the popular Pokémon GO as well as the implementation of the interaction techniques. We conducted a 2×3 within-subjects factorial design study with ten participants to compare the two gestures on a throwing task for three different target sizes. We found significant differences in terms of accuracy, subjective ratings, system usability, and game experiences, but a fifty split for preferences, which we will further discuss in the next chapter under Section 5.2.

Chapter 5

Discussion

From the two user studies, we summarise the key findings and discuss their implications in this chapter. We will cover the discussion of the elicitation study in Section 5.1 and the validation study in Section 5.2. We summarise the answers to our three research questions in Section 5.3 and the implication of this research in Section 5.4. Finally, Section 5.5 will cover the limitations of this research.

5.1 Discussion of Elicitation Study

In the elicitation study, we hypothesised that the elicited surface gestures would be better in terms of *Goodness* and *Ease of Use* (*H1*), and the elicited motion gestures for *Engagement* (*H2*). We found that the results partially supported our hypotheses for some of the tasks. The rating outcomes indicated that the participant's perception of their gestures tied to the nature of the tasks, which involved factors beyond the consideration of the study. Further examination of the results yielded some insights for the tasks that were found to have significant results.

In terms of ***Goodness***, *Surface* gestures were rated significantly higher for Tasks 5 – *Scale-Up*, 9 – *Open Drawer*, and 10 – *Close Drawer*, and *Motion* gestures for Task 7 – *Move*. As we expected, the elicited surface gestures were a familiar legacy set of surface gestures, and the participants found it to be highly suitable, which was reflected in the high average scores for all tasks. Moreover, as mentioned in Subsection 3.6.4, surface gestures were highly *Reusable* and therefore, might be easier to recall. Nevertheless, the motion gestures were also highly rated, particularly, for the moving task, which was done in 3D space. This was an indicator that 3D tasks, which required precise 3D inputs, might benefit from the elicited motion gestures. Hence as pointed out in Subsection 3.6.4 – *Functionality-based* and *Context-based*, it is crucial to consider at the experience-level of each interaction, what are the key factors for a particular experience and which type of gestures would be best to support them.

For ***Ease of Use***, *Surface* gestures were easier to perform for Task 9 – *Open Drawer*, 10 – *Close Drawer*, 11 – *Open Door*, and 12 – *Close Door*. We expected that lower physical demand of the surface gestures made movement tasks easier to perform. For example, in the Open Drawer task, it did not matter how the drawer was opened, whether it was opened slowly or at half the distance, and therefore, any gesture which executed the action would satisfy the goal. However, the results might be different if the same task would be in a gaming context. For example, in stealth games where players have to quietly open the doors or drawers to avoid detection, this would require different types of interaction for a better experience.

For ***Engagement***, *Motion* gestures were rated significantly higher for Task 1 – *Slingshot*, 2 – *Throw*, 7 – *Move*, and 11 – *Open Door*. As predicted and discussed in *Trade-Offs Between Ease of Use and Engagement*, the participants felt that the motion gestures made the interaction more engaging with the mobile AR application. Because the elicited motion gestures reflected the physical movement involved in performing similar physical tasks in the real world. From the overall results, the majority of the participant indicated that they would prefer motion gestures in a gaming context but prefer surface gestures for the other application as it took less effort. From these findings, we propose using *TMR* gestures to improve the level of engagement in mobile AR applications, especially in a gaming context.

5.2 Discussion of Gesture Validation

To validate some of our gestures and recommendations from the elicitation study, we conducted a study to compare the two gestures. We chose to compare *Swipe Up* for *Surface* gesture as a baseline and *Push Forward & Changing Axis* for our proposed *TMR* motion gesture in Task 2 – *Throw* with three target sizes. We hypothesised that there would be significant differences in accuracy (*H1*) subjective ratings (*H2*), game experience (*H3*), and system usability (*H4*). From the study, we found the results to support all of our hypotheses.

In terms of accuracy, we found significant differences between the two interaction techniques across various target sizes. Generally, the *Swipe Up* gesture was more accurate than *Push Forward & Changing Axis,* and the smaller targets were more difficult to hit as expected. We found that for the *Surface* gesture, the participants had better stability and control holding the device in their non-dominant hand and swipe with their dominant hand. One of the participants attempted with a single hand but changed to operate with both hands. Furthermore, the participants could see the screen at all time while interacting, although, their finger might occlude at time. On the contrary, the *Push Forward & Changing Axis* did not allow the participants to see the screen cleary as the device was constantly moving, which took away the stability and control. However, the gesture could be operated single-handed using only the dominant hand.

For the subjective ratings, we found that *Swipe Up* was significantly easier to perform but *Push Forward & Changing Axis* was more engaging, which coincided with the findings from the subjective ratings in the elicitation study. In terms of in-game experience, *Push Forward & Changing Axis* was rated significantly higher than *Swipe Up* on *Competence*, *Sensory and Imaginative Immersion, Flow,* and *Challenge*. This coincided with our findings from Subsection 3.6.4 that *TMR* would benefit the interaction in a gaming context.

TMR made the participants felt challenged to overcome the difficulty of the inaccuracy. At the same time, made them feel the sensory immersion due to the physical and natural movement of their arm to control the device in their palm. The physical activity engaged the participants, making them felt skilful and competence, providing them with an elevated sense of accomplishment when they could successfully achieve their goal. During the interaction using TMR, the participants needed to pay more attention and focus, beyond the increased level of immersion, combining this factor with the sense of challenge, yielded a better sense of flow toward the overall experience.

For the system usability scale (SUS), playing the game using a *Swipe Up* gesture was more usable than *Push Forward & Changing Axis*. We believed that the reason for this was due to a combination of familiarity and ease of use. We observed that some participants could learn to use *TMR* quickly or at least found it novel and interesting to learn, while the others took longer to train, and some felt irritated to use it. This reflected in the number of votes on preference where we received a fifty-fifty split between the two interaction techniques.

5.3 Implications

We have provided the designers with the user-defined surface gesture and motion gesture sets for mobile AR applications. The designers will be able to determine the most suitable gesture for their application from our comprehensive sets of gestures. Furthermore, we have proposed the Touch-Move-Release (TMR) gestures. Note that the user-defined motion gesture set is only a subset of all possible TMR gestures that exhibit TMR design pattern for 3D gesture-based interaction for mobile AR on a handheld device. Any gesture would be considered TMR gestures if it follows the three steps of Touch – to register or initiate the beginning of the action; Move – to manipulate providing the input in 3D space through device movement; and Release – to terminate and commit the change. There are potentially more TMR gestures to be discovered based on the tasks in focus.

5.4 Answers to Research Question

In this research, we asked three research questions. The first research question asked if there exists a gesture-based interaction that can utilise mobile AR three-dimensional space. We have elicited the Touch-Move-Release (TMR) motion gestures and have provided the designers with the user-defined motion gestures for mobile AR applications, as shown in Figure 5. The second research question asked for the user's perception of TMR comparing to the surface gestures. In the same study, we also elicited the user-defined surface gestures for mobile AR. We compared the two sets of gestures in terms of goodness, ease of use, and engagement. We found that, generally, the participants felt that the surface gesture set was easier to use, but the motion gesture set was more engaging. However, the ratings were task- and context-dependent, and we needed to be specific with the type of experiences that we would like to deliver to the user to make the appropriate recommendation. It turned out that TMR was suitable for a gaming application. Finally, the third research question asked how TMR fair against the surface gesture in an actual mobile AR application. We learned that TMR might be suitable for games. Therefore, we developed a Pokémon GO clone and chose a throwing task for the comparison to validate TMR. We found that although TMR was less accurate and more difficult to use than the surface gesture, it was more engaging and offered better in-game experience heighten the level of competence, immersion, flow, and challenge for the player.

5.5 Limitations

There were a number of drawbacks between the two studies that could be addressed. Firstly, although we surveyed current mobile AR applications for suitable tasks, we had to limit the number of selected tasks in the elicitation study to only twelve. This also limits the number of gestures elicited especially for the motion gestures, which has a much larger design space than the surface gestures. Similar studies could be conducted for a specific scenario to elicit a broader but suitable set of gestures to complement our proposed set. Secondly, in the validation study, we could compare only for a single task of the twelve tasks. This was due to the enormous amount of time dedicated to creating game experience replicating the original Pokémon GO. Lastly, we could have measured the task load with questionnaires, e.g. NASA TLX, to compare the cognitive and physical effort in performing the task, however, our focus was on the overall experience from utilising TMR. We could have already predicted indirectly from the SUS result that TMR would likely be considered more demanding for the task load, but this needs to be confirmed.

Beyond the limitations of the studies, we also identified a number of limitations of the proposed TMR gestures. Firstly, there is limited visual feedback due to the movement of the device, which might lead to blurry vision, out of sight, or unsuitable viewing angle of the screen. Secondly, prolonged interaction might lead to fatigue as TMR requires more physical movement. This might not be ideal for experiences with longer duration. Lastly, there is a risk of damaging the device as the user might lose their grip when they move it around.

Chapter 6

Conclusion and Future Work

In this research, we set out to explore novel gesture-based interaction methods for mobile Augmented Reality (AR) applications on handheld devices. From our survey of current mobile AR applications, we found that most applications adopted the conventional surface gestures as the primary input. However, this limited the input to the 2D touchscreen of the handheld devices instead of utilising the 3D space that the AR interface had to offer. This gave rise to our first research question that asked if there might be a gesture-based interaction that could utilise the 3D interaction space of mobile AR. For our second research question, we asked how the 3D gestures would compare to the surface gestures in terms of suitability, ease of use, and engagement.

To answer these two questions, we conducted an elicitation study to elicit two sets of gestures, surface and motion, for twelve tasks selected from six popular AR applications. The study yielded 504 gestures, from which we selected a total of 25 gestures for the final user-defined gesture set, including 13 surface and 12 motion gestures. We compared the two sets of gestures in terms of *Goodness*, *Ease of Use*, and *Engagement*. We found that surface gestures were commonly used conventional set of gestures that participants found significantly easier to use for some tasks, while the motion gestures were significantly more engaging for some other tasks. From these findings, we observed an interaction pattern and proposed the Touch-Move-Release (TMR) interaction technique to improve engagement for mobile AR applications, especially for games.

Following the two research questions, we asked the final question of how would the TMR motion gesture fair against the surface gesture in an actual mobile AR application. To answer this final question, we developed a Pokémon GO clone game with both types of gesture input. We conducted a validation study to compare the two interaction techniques in terms of accuracy with three different target sizes, subjective ratings similar to the previous study, in-game experience, system usability, and preferences. We found that the surface gesture was more accurate and easier to us, but TMR was more engaging and offered better in-game experience heighten the level of competence, immersion, flow, and challenge for the player, which aligned with the results from the elicitation study. We discussed our results and shared the implications of this research as well as pointed out the limitations.

For future work, we would begin by addressing the current limitations of this research. For example, further validation studies would be necessary to verify the validity of the user-defined gesture sets. More tasks should be included in the elicitation study to cover a broader scope of possible AR tasks. There is also an opportunity to use TMR for exergames, i.e. games that have players do more physical activities. Furthermore, we should investigate how to improve accuracy and visual feedback, reduce user fatigue, and lower the risk of damaging the device.

Reference

[1] Daniel Ashbrook and Thad Starner, 2010. MAGIC: a motion gesture design tool. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2159-2168.

[2] Ronald T Azuma, 1997. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments 6*, 4, 355-385.

[3] IKEA Place, https://www.ikea.com/ca/en/customer-service/mobile-apps/(Retrieved on 28/04/2020).

[4] Huidong Bai, 2016. Mobile Augmented Reality: Free-hand Gesture-based Interaction.

[5] Aaron Bangor, Philip Kortum, and James Miller, 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies 4*, 3, 114-123.

[6] Joanna Bergstrom-Lehtovirta and Antti Oulasvirta, 2014. Modeling the functional area of the thumb on mobile touchscreen surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1991-2000.

[7] Mark Billinghurst, Adrian Clark, and Gun Lee, 2015. A survey of augmented reality. *Foundations and Trends® in Human–Computer Interaction 8*, 2-3, 73-272.

[8] Sebastian Boring, Marko Jurmu, and Andreas Butz, 2009. Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7* ACM, 161-168.

[9] Xiang'Anthony' Chen, Tovi Grossman, Daniel J Wigdor, and George Fitzmaurice, 2014. Duet: exploring joint interactions on a smart phone and a smart watch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* ACM, 159-168.

[10] Zhen Chen, Xiaochi Ma, Zeya Peng, Ying Zhou, Mengge Yao, Zheng Ma, Ci Wang, Zaifeng Gao, and Mowei Shen, 2018. User-defined gestures for gestural interaction: extending from hands to other body parts. *International Journal of Human–Computer Interaction 34*, 3, 238-250.

[11] Ashley Colley, Wouter Van Vlaenderen, Johannes Schöning, and Jonna Häkkilä, 2016. Changing the camera-to-screen angle to improve AR browser usage. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* ACM, 442-452.

[12] Nancy J Cooke, 1994. Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies 41*, 6, 801-849.

[13] Klen Čopič Pucihar, Paul Coulton, and Jason Alexander, 2014. The use of surrounding visual context in handheld AR: device vs. user perspective rendering. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* ACM, 197-206.

[14] Microsoft HoloLens 2, https://www.microsoft.com/en-us/hololens/hardware(Retrieved on 28/04/2020).

[15] Jian Cui, Arjan Kuijper, Dieter W Fellner, and Alexei Sourin, 2016. Understanding people's mental models of mid-air interaction for virtual assembly and shape modeling. In *Proceedings of the 29th International Conference on Computer Animation and Social Agents*, 139-146.

[16] Linda Di Geronimo, Marica Bertarini, Julia Badertscher, Maria Husmann, and Moira C Norrie, 2017. Exploiting mid-air gestures to share data among devices. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 1-11.

[17] Julien Epps, Serge Lichman, and Mike Wu, 2006. A study of hand shape use in tabletop gesture interaction. In *CHI'06 extended abstracts on human factors in computing systems*, 748-753.

[18] George H. Forman and John Zahorjan, 1994. The challenges of mobile computing. *Computer 27*, 4, 38-47.

[19] Mayank Goel, Jacob Wobbrock, and Shwetak Patel, 2012. GripSense: using built-in sensors to detect hand posture and pressure on commodity mobile phones. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, 545-554.

[20] Eg Su Goh, Mohd Shahrizal Sunar, and Ajune Wanis Ismail, 2019. 3D Object Manipulation Techniques in Handheld Mobile Augmented Reality Interface: A Review. *IEEE Access 7*, 40581-40601.

[21] Daniel Goldsmith, Fotis Liarokapis, Garry Malone, and John Kemp, 2008. Augmented reality environmental monitoring using wireless sensor networks. In *2008 12th International Conference Information Visualisation* IEEE, 539-544.

[22] Taejin Ha and Woontack Woo, 2011. ARWand: Phone-based 3d object manipulation in augmented reality environment. In *2011 International Symposium on Ubiquitous Virtual Reality* IEEE, 44-47.

[23] Björn Hartmann, Leith Abdulla, Manas Mittal, and Scott R Klemmer, 2007. Authoring sensor-based interactions by demonstration with direct manipulation and pattern recognition. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 145-154.

[24] Hayati Havlucu, Mehmet Yarkın Ergin, Idil Bostan, Oğuz Turan Buruk, Tilbe Göksun, and Oğuzhan Özcan, 2017. It made more sense: Comparison of user-elicited on-skin touch and freehand gesture sets. In *International Conference on Distributed, Ambient, and Pervasive Interactions* Springer, 159-171.

[25] Anders Henrysson, Mark Billinghurst, and Mark Ollila, 2005. Face to face collaborative AR on mobile phones. In *Fourth ieee and acm international symposium on mixed and augmented reality (ismar'05)* IEEE, 80-89.

[26] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani, 2014. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* ACM, 1063-1072.

[27] Ken Hinckley, Jeff Pierce, Mike Sinclair, and Eric Horvitz, 2000. Sensing techniques for mobile interaction. In *Proceedings of the 13th annual ACM symposium on User interface software and technology*, 91-100.

[28] Wolfgang Hürst and Casper Van Wezel, 2013. Gesture-based interaction via finger tracking for mobile augmented reality. *Multimedia Tools and Applications 62*, 1, 233-258.

[29] WA IJsselsteijn, YAW De Kort, and Karolien Poels, 2013. The game experience questionnaire. *Eindhoven: Technische Universiteit Eindhoven*.

[30] ARKit 3.5, https://developer.apple.com/augmented-reality/(Retrieved on 28/04/2020).

[31] ARCore, https://developers.google.com/ar/discover(Retrieved on 28/04/2020).

[32] Eleanor Jones, Jason Alexander, Andreas Andreou, Pourang Irani, and Sriram Subramanian, 2010. GesText: accelerometer-based gestural text-entry systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2173-2182.

[33] Stan Kurkovsky, Ranjana Koshy, Vivian Novak, and Peter Szul, 2012. Current issues in handheld augmented reality. In *2012 international conference on communications and information technology (ICCIT)* IEEE, 68-72.

[34] Victor Kyriazakos, Giorgos Nikolakis, and Konstantinos Moustakas, 2016. Natural Interaction with 3D Content on Mobile AR Systems Using Gesture Recognition. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics* Springer, 348-357.

[35] Gun A Lee, Ungyeon Yang, Yongwan Kim, Dongsik Jo, Ki-Hong Kim, Jae Ha Kim, and Jin Sung Choi, 2009. Freeze-Set-Go interaction method for handheld mobile augmented reality environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, 143-146.

[36] Keenan R May, Thomas M Gable, and Bruce N Walker, 2017. Designing an in-vehicle air gesture set using elicitation methods. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 74-83.

[37] Meredith Ringel Morris, Jacob O Wobbrock, and Andrew D Wilson, 2010. Understanding users' preferences for surface gestures. In *Proceedings of graphics interface 2010* Canadian Information Processing Society, 261-268.

[38] Annette Mossel, Benjamin Venditti, and Hannes Kaufmann, 2013. 3DTouch and HOMER-S: intuitive manipulation techniques for one-handed handheld augmented reality. In *Proceedings of the Virtual Reality International Conference: Laval Virtual*, 1-10.

[39] Pokemon Go, https://pokemongolive.com/en/(Retrieved on 28/04/2020).

[40] Tran Pham, Jo Vermeulen, Anthony Tang, and Lindsay MacDonald Vermeulen, 2018. Scale impacts elicited gestures for manipulating holograms: Implications for AR gesture design. In *Proceedings of the 2018 Designing Interactive Systems Conference*, 227-240.

[41] Thammathip Piumsomboon, David Altimira, Hyungon Kim, Adrian Clark, Gun Lee, and Mark Billinghurst, 2014. Grasp-Shell vs gesture-speech: A comparison of direct and indirect natural interaction techniques in augmented reality. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* IEEE, pp.73-82. IEEE,2014.

[42] Thammathip Piumsomboon, Adrian Clark, and Mark Billinghurst, 2013. KITE: Platform for mobile Augmented Reality gaming and interaction using magnetic

tracking and depth sensing. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* IEEE, pp. 289-290. IEEE,2013.

[43] Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn, 2013. User-defined gestures for augmented reality. In *IFIP Conference on Human-Computer Interaction* Springer, 282-299.

[44] PTC Software, https://www.ptc.com/en(Retrieved on 28/04/2020).

[45] Vuforia Engine, https://engine.vuforia.com/engine(Retrieved on 28/04/2020).

[46] Qualcomm Semiconductor company, https://www.qualcomm.com/(Retrieved on 28/04/2020).

[47] QuiverVision, http://www.quivervision.com(Retrieved on 28/04/2020).

[48] Julie Rico and Stephen Brewster, 2010. Usable gestures for mobile interfaces: evaluating social acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* ACM, 887-896.

[49] Jaime Ruiz, Yang Li, and Edward Lank, 2011. User-defined motion gestures for mobile interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* ACM, 197-206.

[50] Jochen Schiller and Agnès Voisard, 2004. *Location-based services*. Elsevier.

[51] UAB. SketchAR, 2019. Start Drawing Easily Using Augmented Reality. SketchAR SketchAR, UAB., sketchar.tech.

[52] Darko Stanimirovic and Daniel Kurz, 2014. Smartwatch-aided handheld augmented reality. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* IEEE, 373-374.

[53] Tomohiro Tanikawa, Hidenori Uzuka, Takuji Narumi, and Michitaka Hirose, 2015. Integrated view-input AR interaction for virtual object manipulation using tablets and smartphones. In *Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology* ACM, 7.

[54] Unity Engine, https://unity.com/(Retrieved on 28/04/2020).

[55] DWF Van Krevelen and Ronald Poelman, 2010. A survey of augmented reality technologies, applications and limitations. *International journal of virtual reality 9*, 2, 1-20.

[56] Stephen Voida, Mark Podlaseck, Rick Kjeldsen, and Claudio Pinhanez, 2005. A study on the manipulation of 2D objects in a projector/camera-based augmented reality environment. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 611-620.

[57] Robert Walter, Gilles Bailly, Nina Valkanova, and Jörg Müller, 2014. Cuenesics: using mid-air gestures to select items on interactive public displays. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*, 299-308.

[58] Daniel Wigdor and Ravin Balakrishnan, 2003. TiltText: using tilt for text input to mobile phones. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*, 81-90.

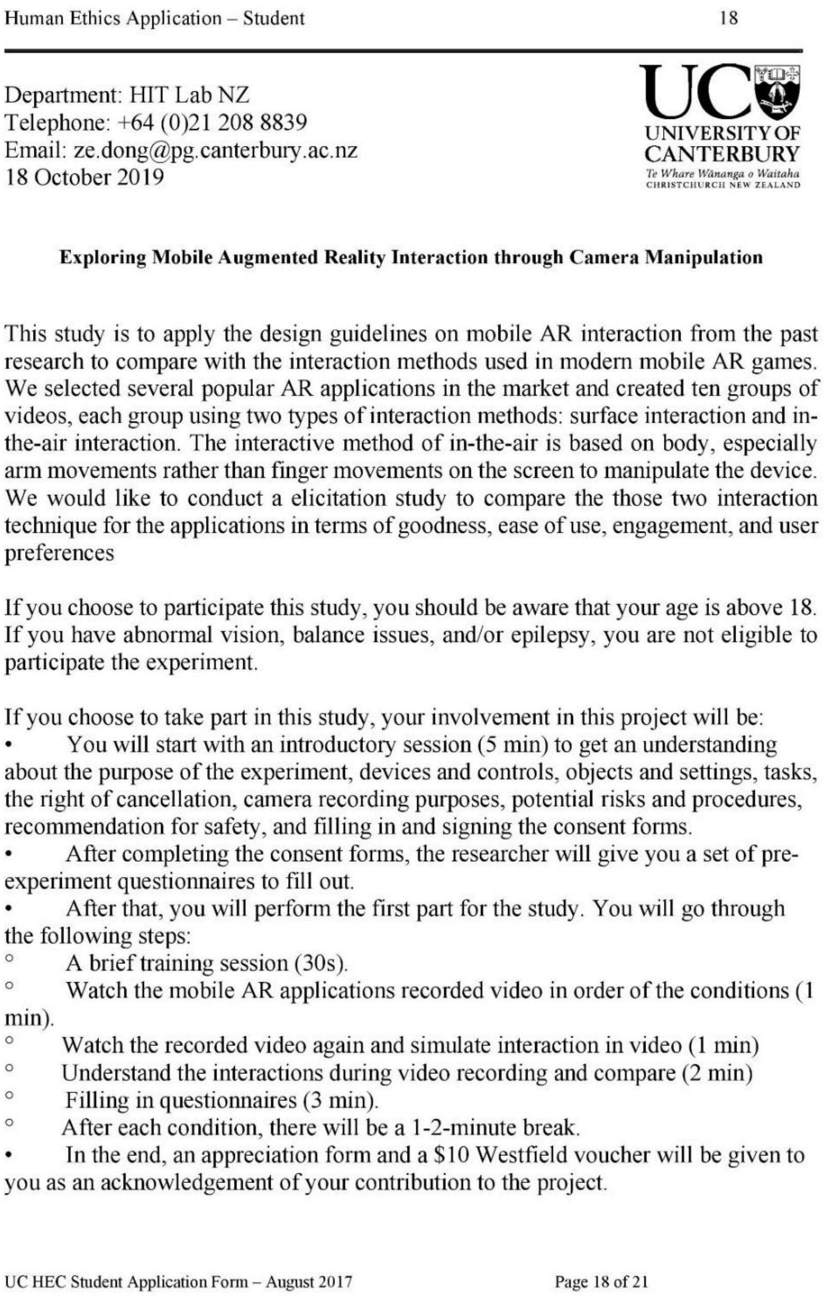
[59] Andrew D Wilson, Shahram Izadi, Otmar Hilliges, Armando Garcia-Mendoza, and David Kirk, 2008. Bringing physics to the surface. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, 67-76.

[60] Jacob O Wobbrock, Meredith Ringel Morris, and Andrew D Wilson, 2009. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* ACM, 1083-1092.

[61] Mike Wu, Chia Shen, Kathy Ryall, Clifton Forlines, and Ravin Balakrishnan, 2006. Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'06)* IEEE, 8 pp.

Appendix A

A.1 Information Sheet



Human Ethics Application – Student 18

Department: HIT Lab NZ
Telephone: +64 (0)21 208 8839
Email: ze.dong@pg.canterbury.ac.nz
18 October 2019

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Exploring Mobile Augmented Reality Interaction through Camera Manipulation

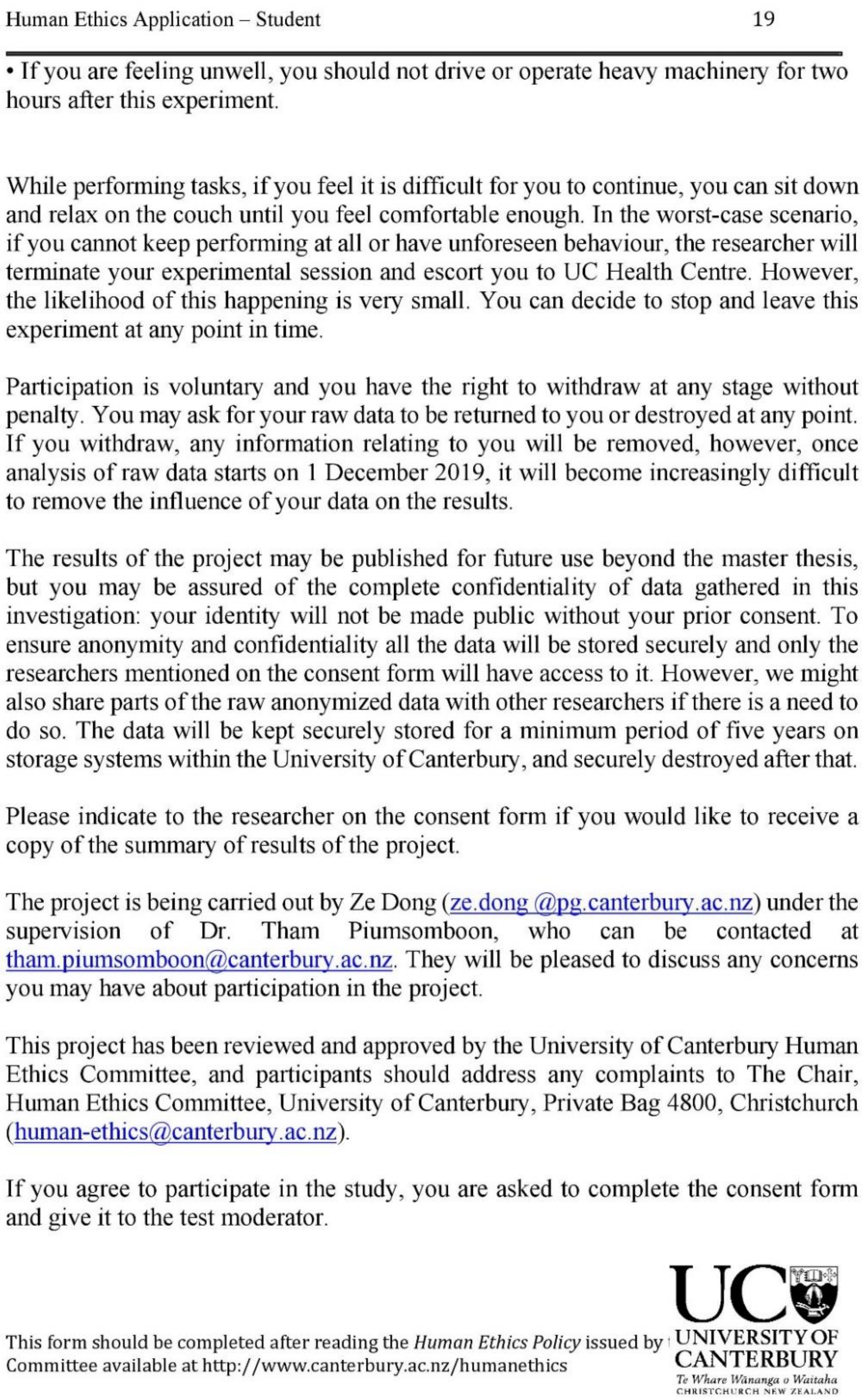
This study is to apply the design guidelines on mobile AR interaction from the past research to compare with the interaction methods used in modern mobile AR games. We selected several popular AR applications in the market and created ten groups of videos, each group using two types of interaction methods: surface interaction and in-the-air interaction. The interactive method of in-the-air is based on body, especially arm movements rather than finger movements on the screen to manipulate the device. We would like to conduct a elicitation study to compare the those two interaction technique for the applications in terms of goodness, ease of use, engagement, and user preferences

If you choose to participate this study, you should be aware that your age is above 18. If you have abnormal vision, balance issues, and/or epilepsy, you are not eligible to participate the experiment.

If you choose to take part in this study, your involvement in this project will be:

- You will start with an introductory session (5 min) to get an understanding about the purpose of the experiment, devices and controls, objects and settings, tasks, the right of cancellation, camera recording purposes, potential risks and procedures, recommendation for safety, and filling in and signing the consent forms.
- After completing the consent forms, the researcher will give you a set of pre-experiment questionnaires to fill out.
- After that, you will perform the first part for the study. You will go through the following steps:
  - A brief training session (30s).
  - Watch the mobile AR applications recorded video in order of the conditions (1 min).
  - Watch the recorded video again and simulate interaction in video (1 min)
  - Understand the interactions during video recording and compare (2 min)
  - Filling in questionnaires (3 min).
  - After each condition, there will be a 1-2-minute break.
- In the end, an appreciation form and a $10 Westfield voucher will be given to you as an acknowledgement of your contribution to the project.

UC HEC Student Application Form – August 2017 Page 18 of 21



• If you are feeling unwell, you should not drive or operate heavy machinery for two hours after this experiment.

While performing tasks, if you feel it is difficult for you to continue, you can sit down and relax on the couch until you feel comfortable enough. In the worst-case scenario, if you cannot keep performing at all or have unforeseen behaviour, the researcher will terminate your experimental session and escort you to UC Health Centre. However, the likelihood of this happening is very small. You can decide to stop and leave this experiment at any point in time.

Participation is voluntary and you have the right to withdraw at any stage without penalty. You may ask for your raw data to be returned to you or destroyed at any point. If you withdraw, any information relating to you will be removed, however, once analysis of raw data starts on 1 December 2019, it will become increasingly difficult to remove the influence of your data on the results.

The results of the project may be published for future use beyond the master thesis, but you may be assured of the complete confidentiality of data gathered in this investigation: your identity will not be made public without your prior consent. To ensure anonymity and confidentiality all the data will be stored securely and only the researchers mentioned on the consent form will have access to it. However, we might also share parts of the raw anonymized data with other researchers if there is a need to do so. The data will be kept securely stored for a minimum period of five years on storage systems within the University of Canterbury, and securely destroyed after that.

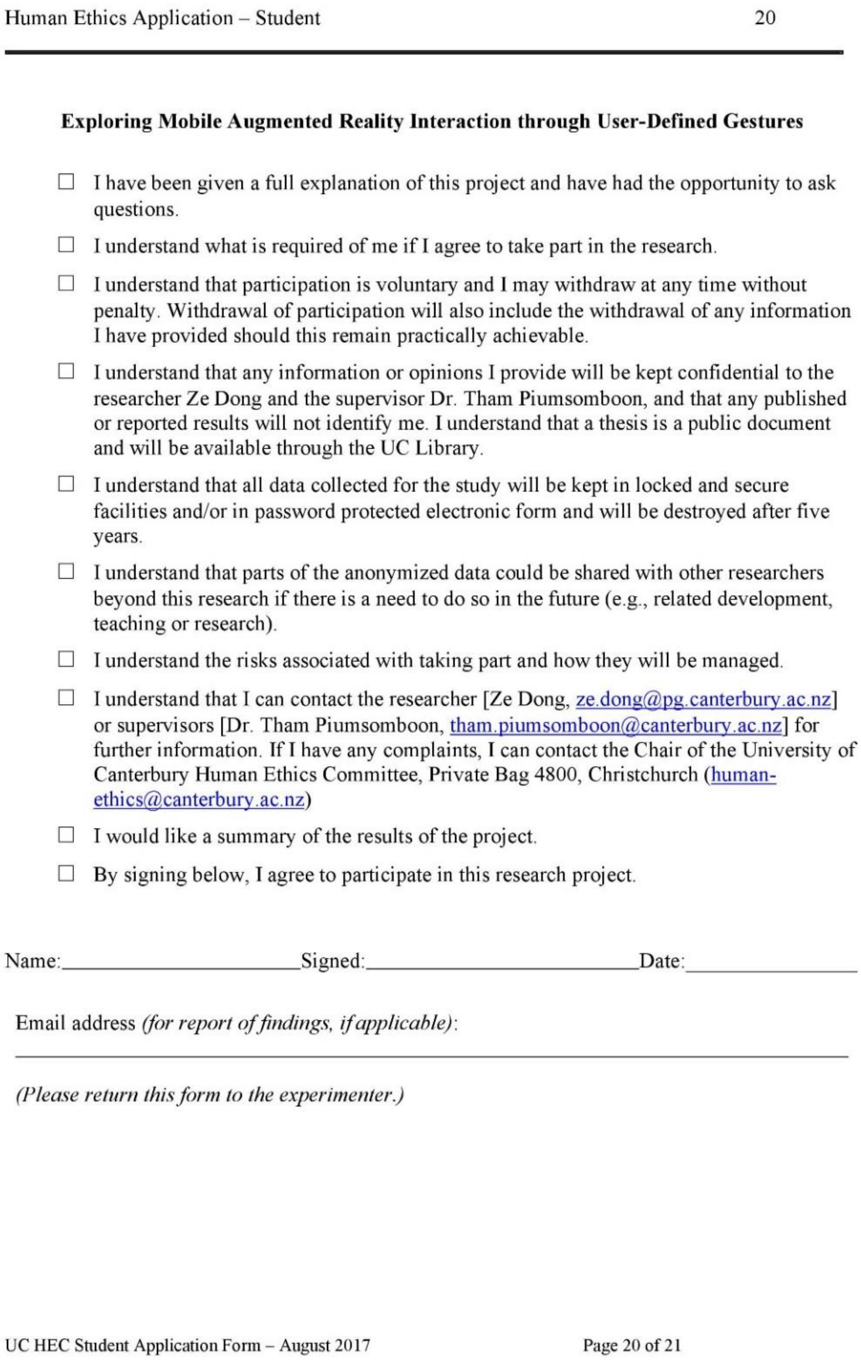
Please indicate to the researcher on the consent form if you would like to receive a copy of the summary of results of the project.

The project is being carried out by Ze Dong (ze.dong @pg.canterbury.ac.nz) under the supervision of Dr. Tham Piumsomboon, who can be contacted at tham.piumsomboon@canterbury.ac.nz. They will be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

If you agree to participate in the study, you are asked to complete the consent form and give it to the test moderator.

A.2 Consent Form



Exploring Mobile Augmented Reality Interaction through User-Defined Gestures

- ☐ I have been given a full explanation of this project and have had the opportunity to ask questions.
- ☐ I understand what is required of me if I agree to take part in the research.
- ☐ I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.
- ☐ I understand that any information or opinions I provide will be kept confidential to the researcher Ze Dong and the supervisor Dr. Tham Piumsomboon, and that any published or reported results will not identify me. I understand that a thesis is a public document and will be available through the UC Library.
- ☐ I understand that all data collected for the study will be kept in locked and secure facilities and/or in password protected electronic form and will be destroyed after five years.
- ☐ I understand that parts of the anonymized data could be shared with other researchers beyond this research if there is a need to do so in the future (e.g., related development, teaching or research).
- ☐ I understand the risks associated with taking part and how they will be managed.
- ☐ I understand that I can contact the researcher [Ze Dong, ze.dong@pg.canterbury.ac.nz] or supervisors [Dr. Tham Piumsomboon, tham.piumsomboon@canterbury.ac.nz] for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz)
- ☐ I would like a summary of the results of the project.
- ☐ By signing below, I agree to participate in this research project.

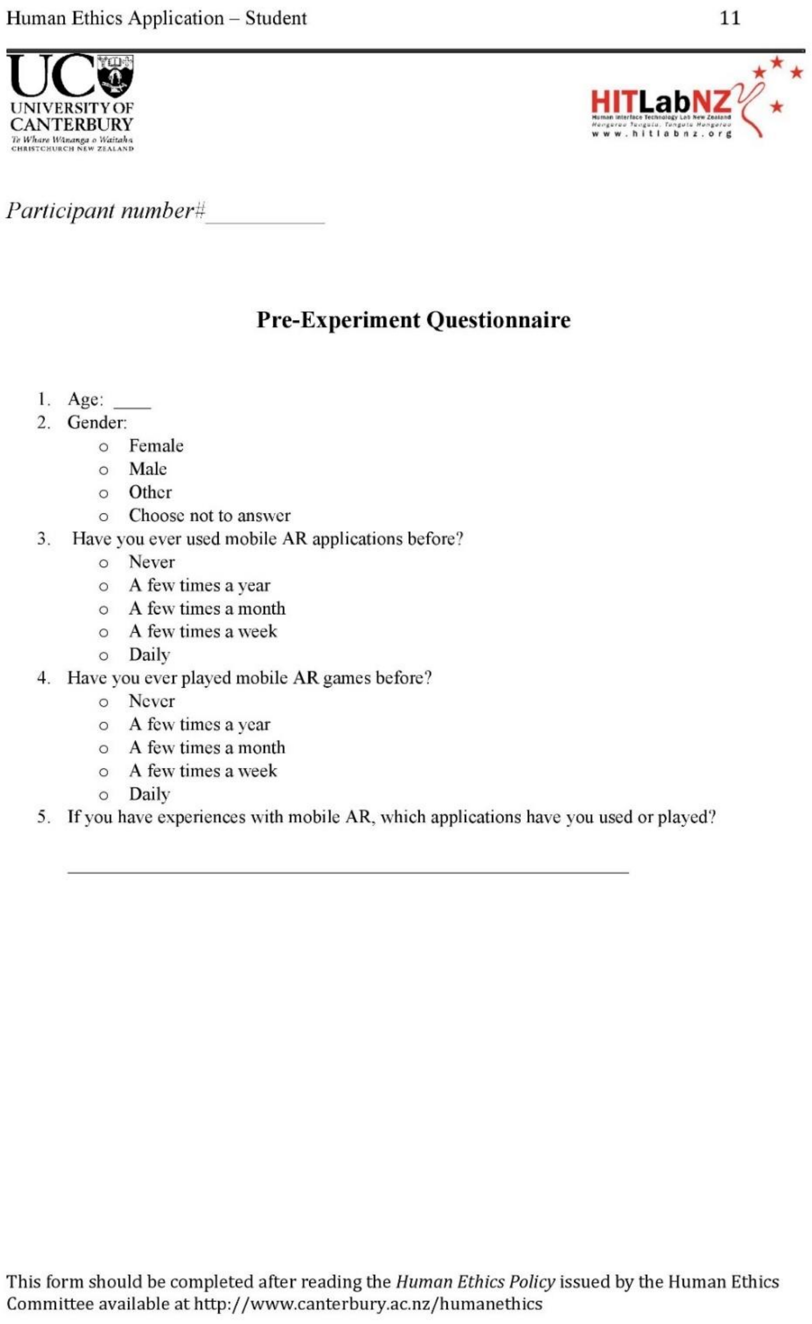
Name:______________________Signed:________________________Date:________________

Email address *(for report of findings, if applicable)*:

____

(Please return this form to the experimenter.)

A.3 Pre-Experiment Questionnaire



Human Ethics Application – Student 11

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HITLabNZ
www.hitlabnz.org

*Participant number#*__________

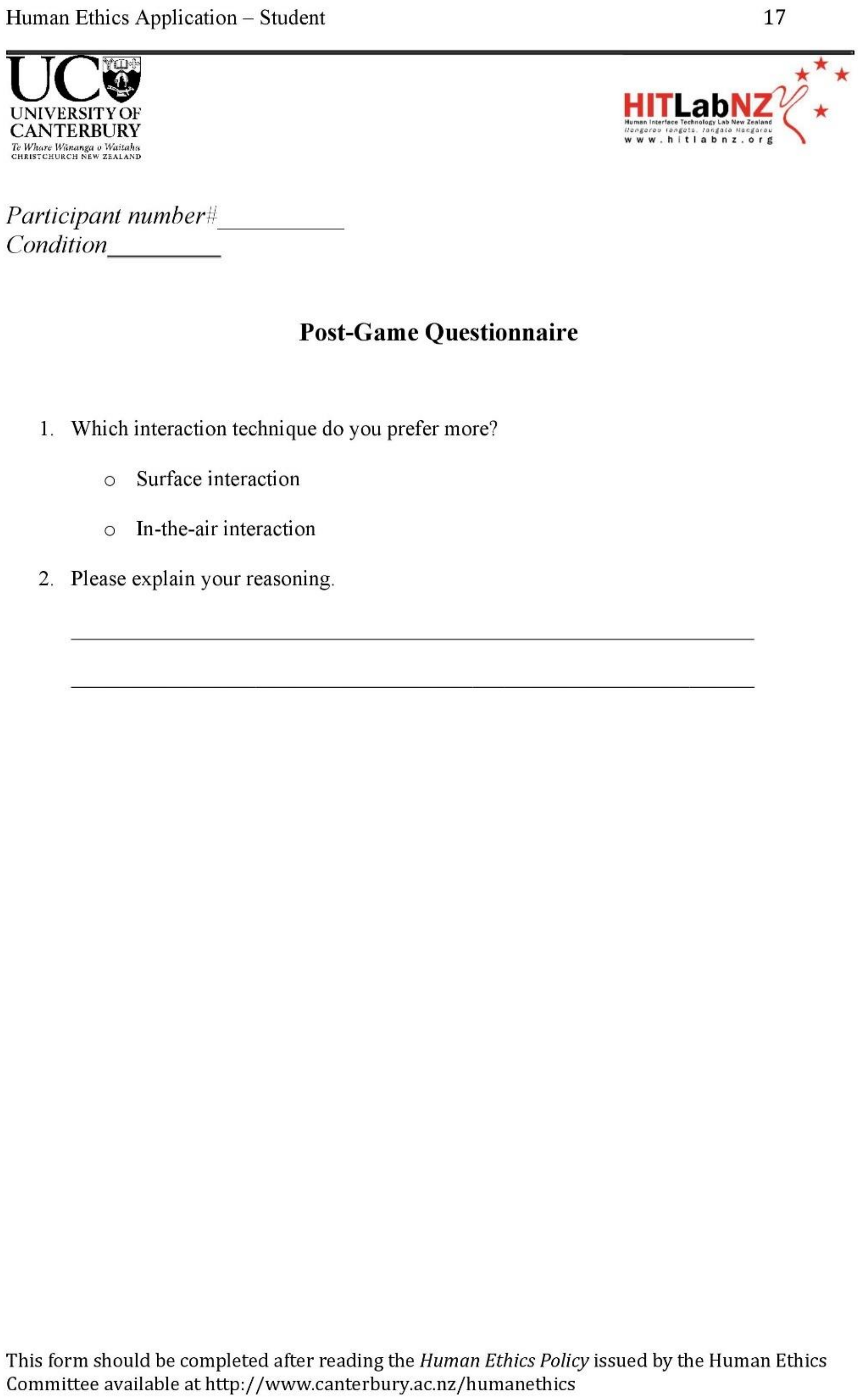
Pre-Experiment Questionnaire

1. Age: ____
2. Gender:
    - ○ Female
    - ○ Male
    - ○ Other
    - ○ Choose not to answer
3. Have you ever used mobile AR applications before?
    - ○ Never
    - ○ A few times a year
    - ○ A few times a month
    - ○ A few times a week
    - ○ Daily
4. Have you ever played mobile AR games before?
    - ○ Never
    - ○ A few times a year
    - ○ A few times a month
    - ○ A few times a week
    - ○ Daily
5. If you have experiences with mobile AR, which applications have you used or played?

    ____

This form should be completed after reading the *Human Ethics Policy* issued by the Human Ethics Committee available at http://www.canterbury.ac.nz/humanethics

A.4 Post-Experiment Questionnaire



Human Ethics Application – Student 17

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HITLabNZ
Human Interface Technology Lab New Zealand
www.hitlabnz.org

*Participant number#*__________
*Condition*_________

Post-Game Questionnaire

1. Which interaction technique do you prefer more?
   - ○ Surface interaction
   - ○ In-the-air interaction
2. Please explain your reasoning.

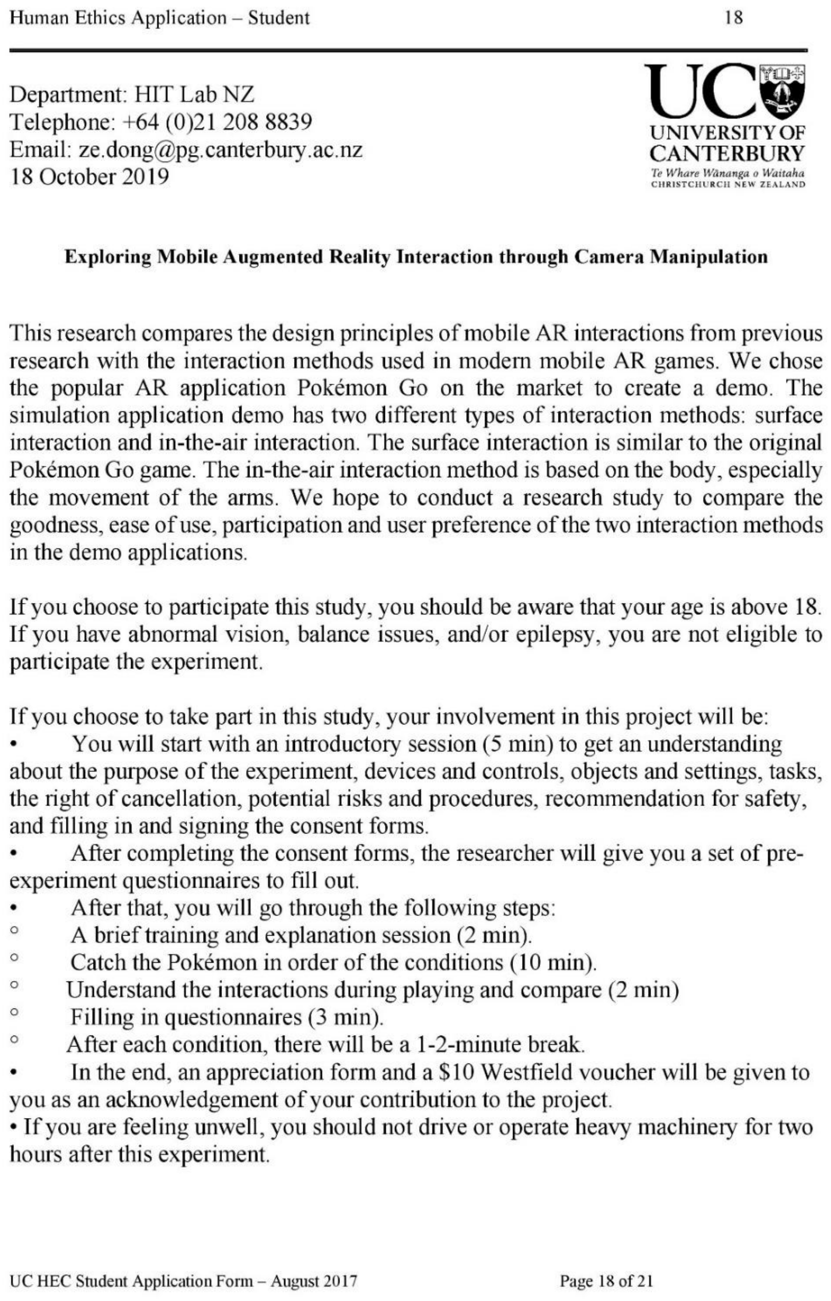
   ____

   ____

This form should be completed after reading the *Human Ethics Policy* issued by the Human Ethics Committee available at http://www.canterbury.ac.nz/humanethics

Appendix B

B.1 Information Sheet



Human Ethics Application – Student 18

Department: HIT Lab NZ
Telephone: +64 (0)21 208 8839
Email: ze.dong@pg.canterbury.ac.nz
18 October 2019

UC
UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Exploring Mobile Augmented Reality Interaction through Camera Manipulation

This research compares the design principles of mobile AR interactions from previous research with the interaction methods used in modern mobile AR games. We chose the popular AR application Pokémon Go on the market to create a demo. The simulation application demo has two different types of interaction methods: surface interaction and in-the-air interaction. The surface interaction is similar to the original Pokémon Go game. The in-the-air interaction method is based on the body, especially the movement of the arms. We hope to conduct a research study to compare the goodness, ease of use, participation and user preference of the two interaction methods in the demo applications.

If you choose to participate this study, you should be aware that your age is above 18. If you have abnormal vision, balance issues, and/or epilepsy, you are not eligible to participate the experiment.

If you choose to take part in this study, your involvement in this project will be:

- You will start with an introductory session (5 min) to get an understanding about the purpose of the experiment, devices and controls, objects and settings, tasks, the right of cancellation, potential risks and procedures, recommendation for safety, and filling in and signing the consent forms.
- After completing the consent forms, the researcher will give you a set of pre-experiment questionnaires to fill out.
- After that, you will go through the following steps:
  - A brief training and explanation session (2 min).
  - Catch the Pokémon in order of the conditions (10 min).
  - Understand the interactions during playing and compare (2 min)
  - Filling in questionnaires (3 min).
  - After each condition, there will be a 1-2-minute break.
- In the end, an appreciation form and a $10 Westfield voucher will be given to you as an acknowledgement of your contribution to the project.
- If you are feeling unwell, you should not drive or operate heavy machinery for two hours after this experiment.

UC HEC Student Application Form – August 2017 Page 18 of 21

While performing tasks, if you feel it is difficult for you to continue, you can sit down and relax on the couch until you feel comfortable enough. In the worst-case scenario, if you cannot keep performing at all or have unforeseen behaviour, the researcher will terminate your experimental session and escort you to UC Health Centre. However, the likelihood of this happening is very small. You can decide to stop and leave this experiment at any point in time.

Participation is voluntary and you have the right to withdraw at any stage without penalty. You may ask for your raw data to be returned to you or destroyed at any point. If you withdraw, any information relating to you will be removed, however, once analysis of raw data starts on 1 December 2019, it will become increasingly difficult to remove the influence of your data on the results.

The results of the project may be published for future use beyond the master thesis, but you may be assured of the complete confidentiality of data gathered in this investigation: your identity will not be made public without your prior consent. To ensure anonymity and confidentiality all the data will be stored securely and only the researchers mentioned on the consent form will have access to it. However, we might also share parts of the raw anonymized data with other researchers if there is a need to do so. The data will be kept securely stored for a minimum period of five years on storage systems within the University of Canterbury, and securely destroyed after that.

Please indicate to the researcher on the consent form if you would like to receive a copy of the summary of results of the project.

The project is being carried out by Ze Dong (ze.dong @pg.canterbury.ac.nz) under the supervision of Dr. Tham Piumsomboon, who can be contacted at tham.piumsomboon@canterbury.ac.nz. They will be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

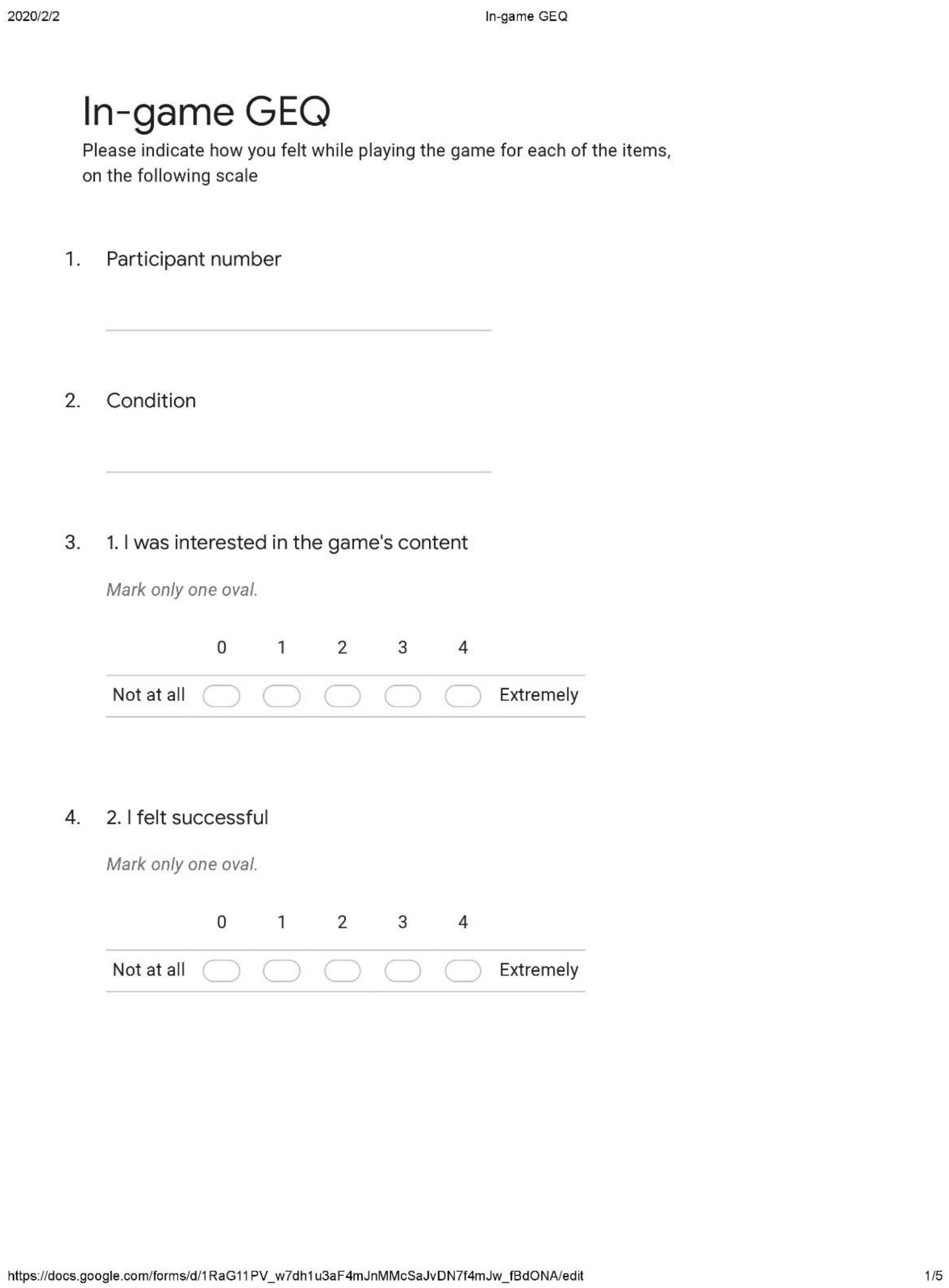
If you agree to participate in the study, you are asked to complete the consent form and give it to the test moderator.

Department: HIT Lab NZ
Telephone: +64 (0)21 208 8839
Email: ze.dong@pg.canterbury.ac.nz
18 October 2019

UC
UNIVERSITY OF
CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

This form should be completed after reading the *Human Ethics Policy* issued by the Human Ethics Committee available at http://www.canterbury.ac.nz/humanethics

B.2 In-Game Experience Questionnaire



2020/2/2 In-game GEQ

In-game GEQ

Please indicate how you felt while playing the game for each of the items, on the following scale

1. Participant number

____

2. Condition

____

3. 1. I was interested in the game's content

Mark only one oval.

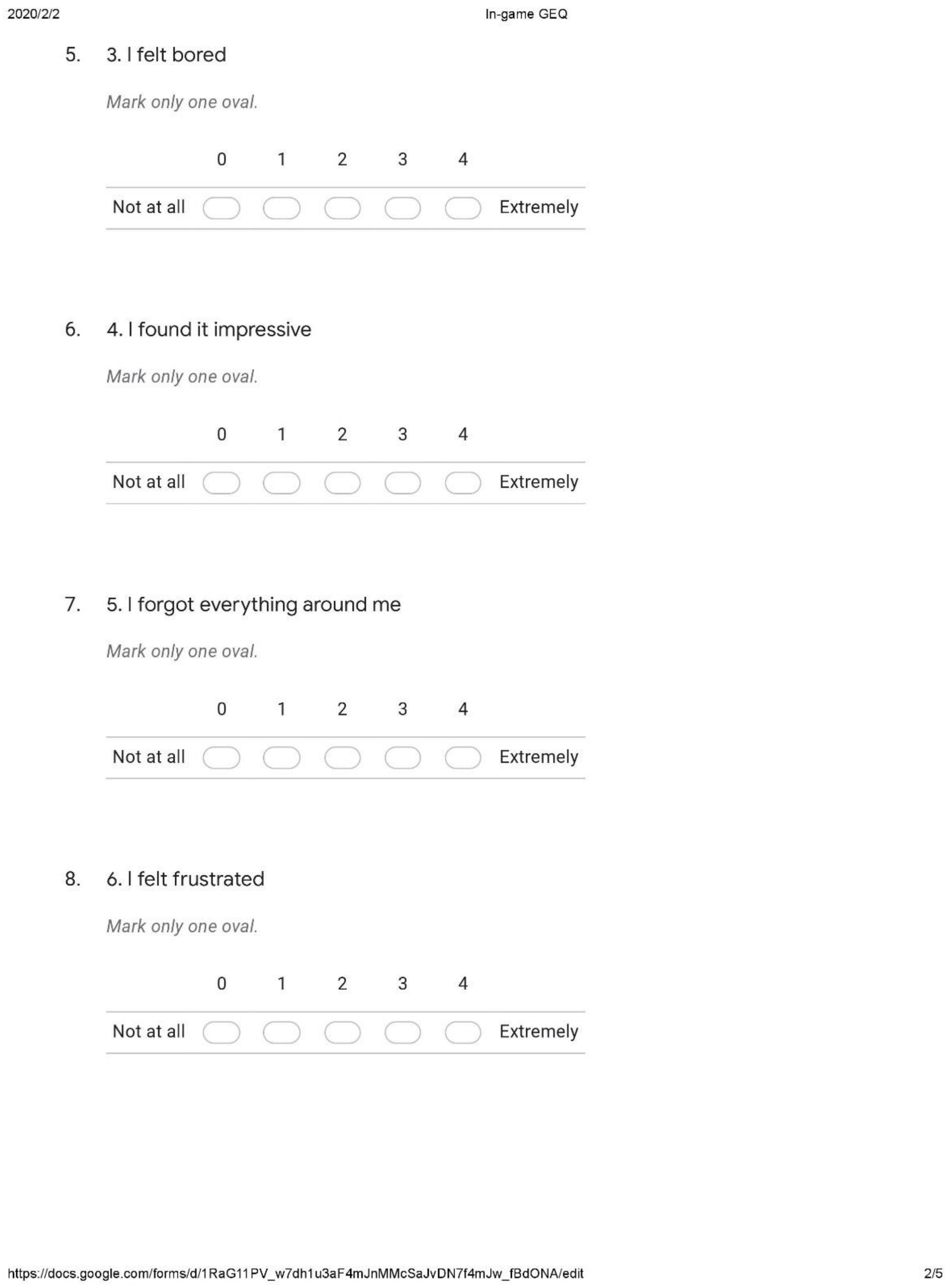
	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

4. 2. I felt successful

Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

https://docs.google.com/forms/d/1RaG11PV_w7dh1u3aF4mJnMMcSaJvDN7f4mJw_fBdONA/edit 1/5



5. 3. I felt bored

Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

6. 4. I found it impressive

Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

7. 5. I forgot everything around me

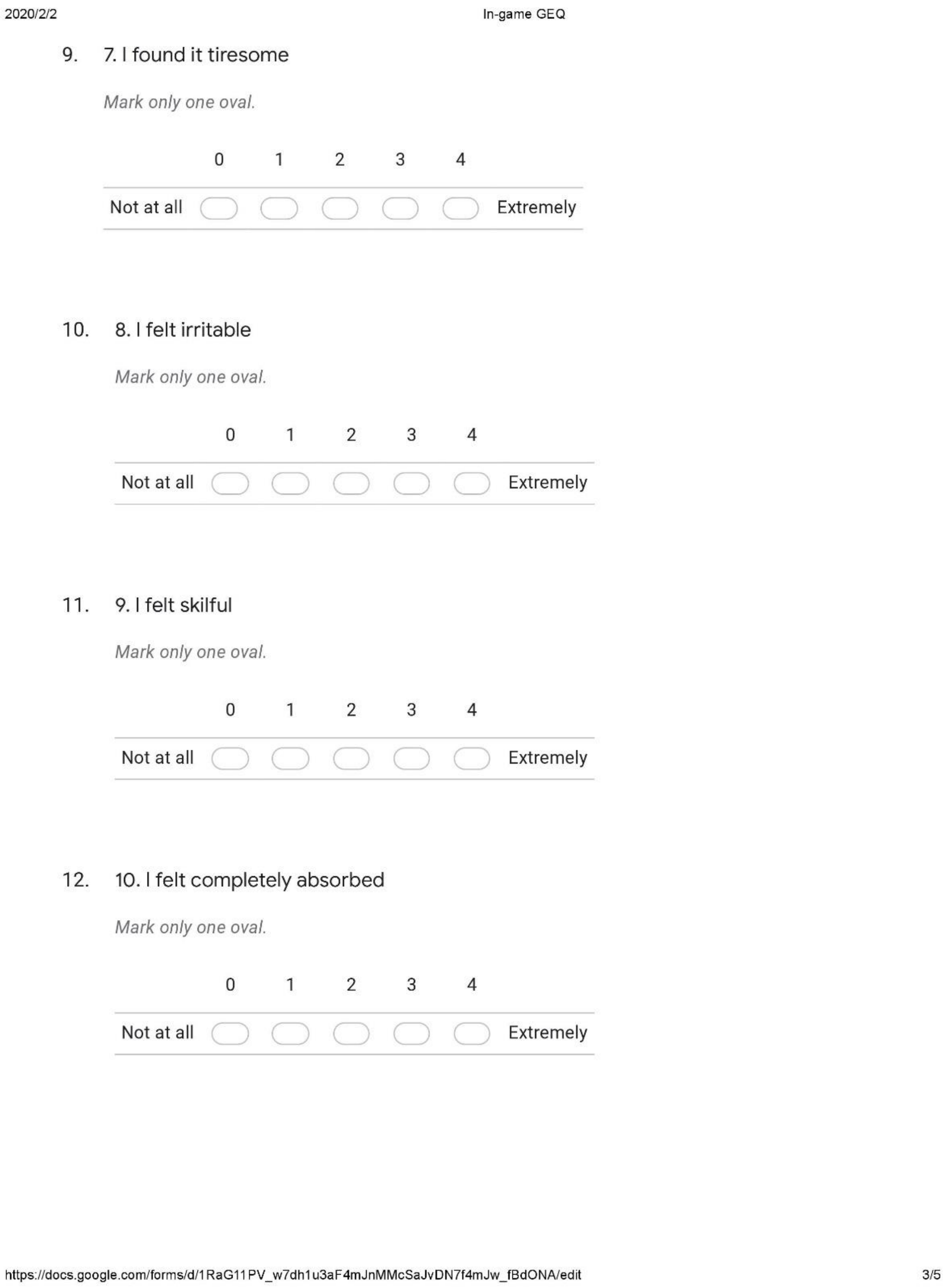
Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

8. 6. I felt frustrated

Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely



2020/2/2 In-game GEQ

9. 7. I found it tiresome

Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

10. 8. I felt irritable

Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

11. 9. I felt skilful

Mark only one oval.

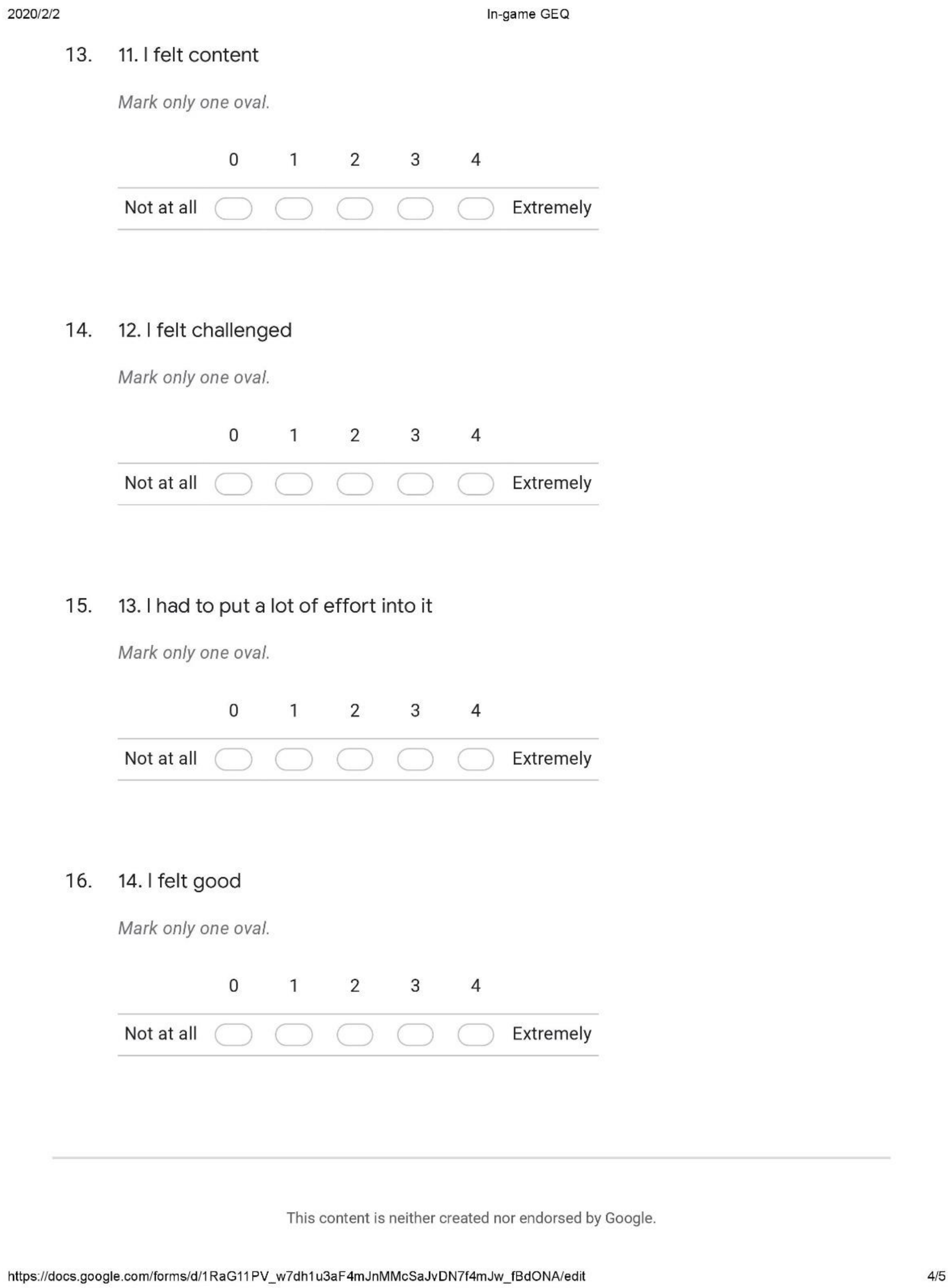
	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

12. 10. I felt completely absorbed

Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

https://docs.google.com/forms/d/1RaG11PV_w7dh1u3aF4mJnMMcSaJvDN7f4mJw_fBdONA/edit 3/5



13. 11. I felt content

Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

14. 12. I felt challenged

Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

15. 13. I had to put a lot of effort into it

Mark only one oval.

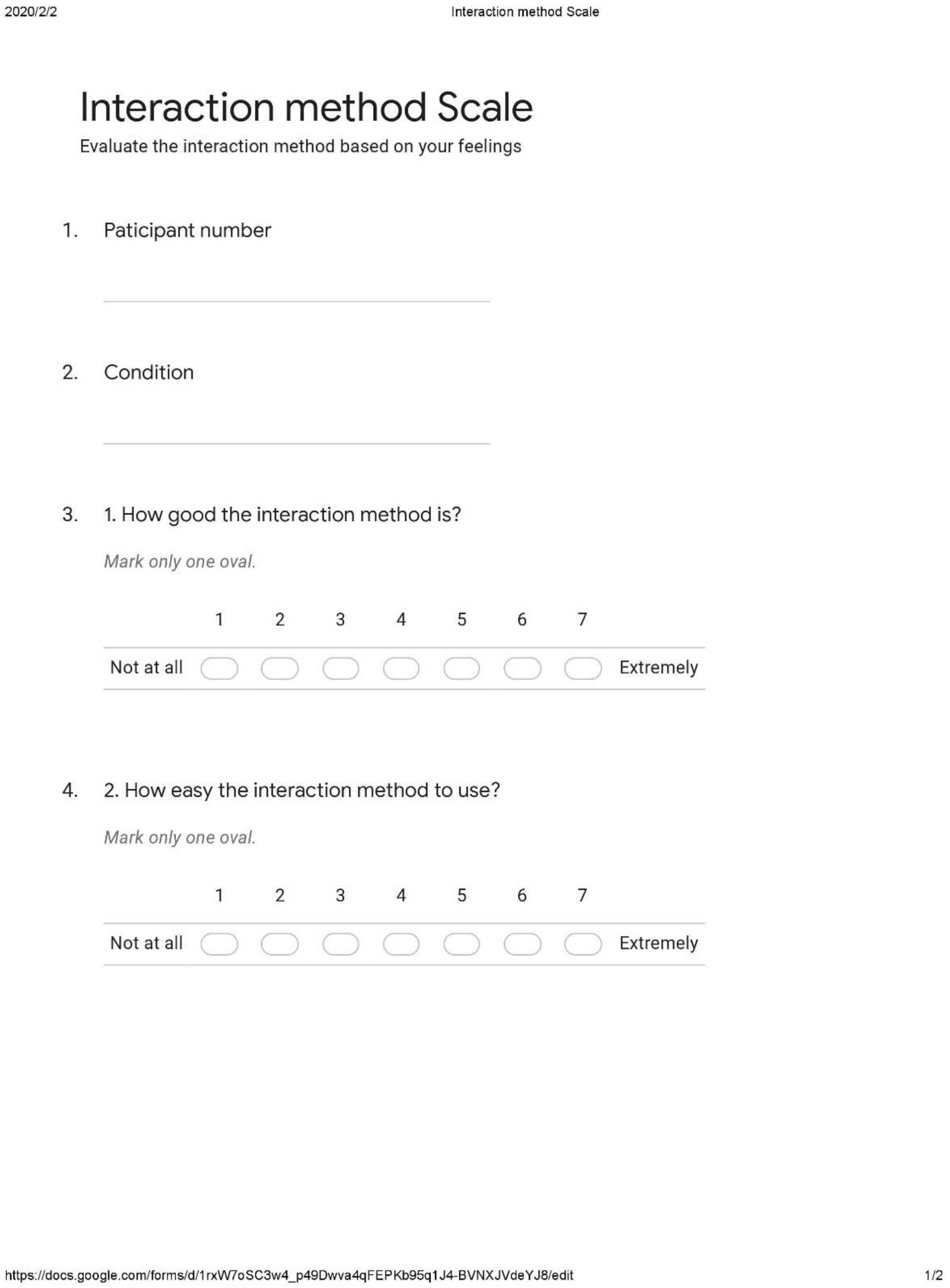
	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

16. 14. I felt good

Mark only one oval.

	0	1	2	3	4	
Not at all	○	○	○	○	○	Extremely

B.3 Interaction Method Scale



2020/2/2 Interaction method Scale

Interaction method Scale

Evaluate the interaction method based on your feelings

1. Paticipant number

2. Condition

3. 1. How good the interaction method is?

Mark only one oval.

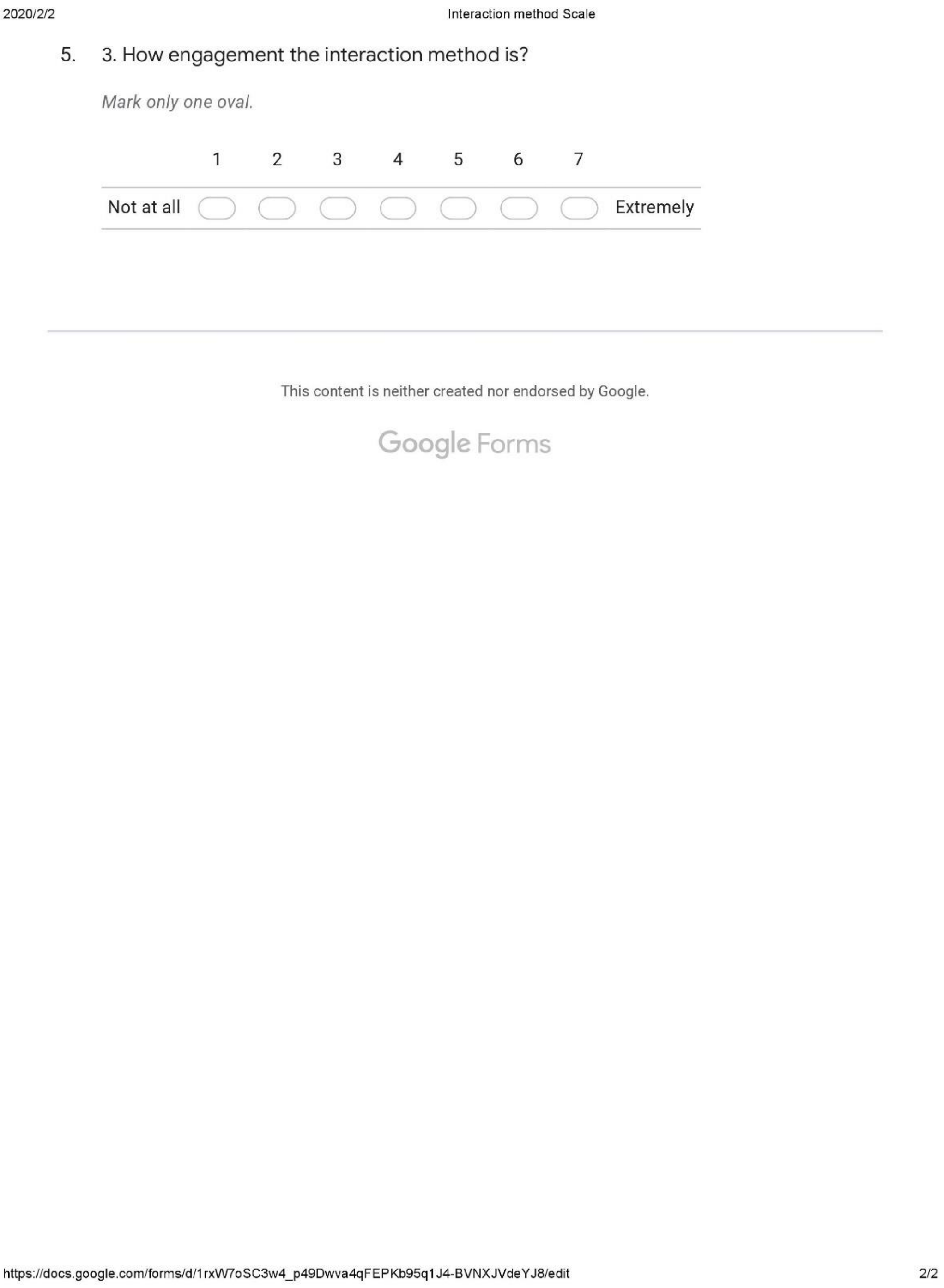
	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Extremely

4. 2. How easy the interaction method to use?

Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Extremely

https://docs.google.com/forms/d/1rxW7oSC3w4_p49Dwva4qFEPKb95q1J4-BVNXJVdeYJ8/edit 1/2

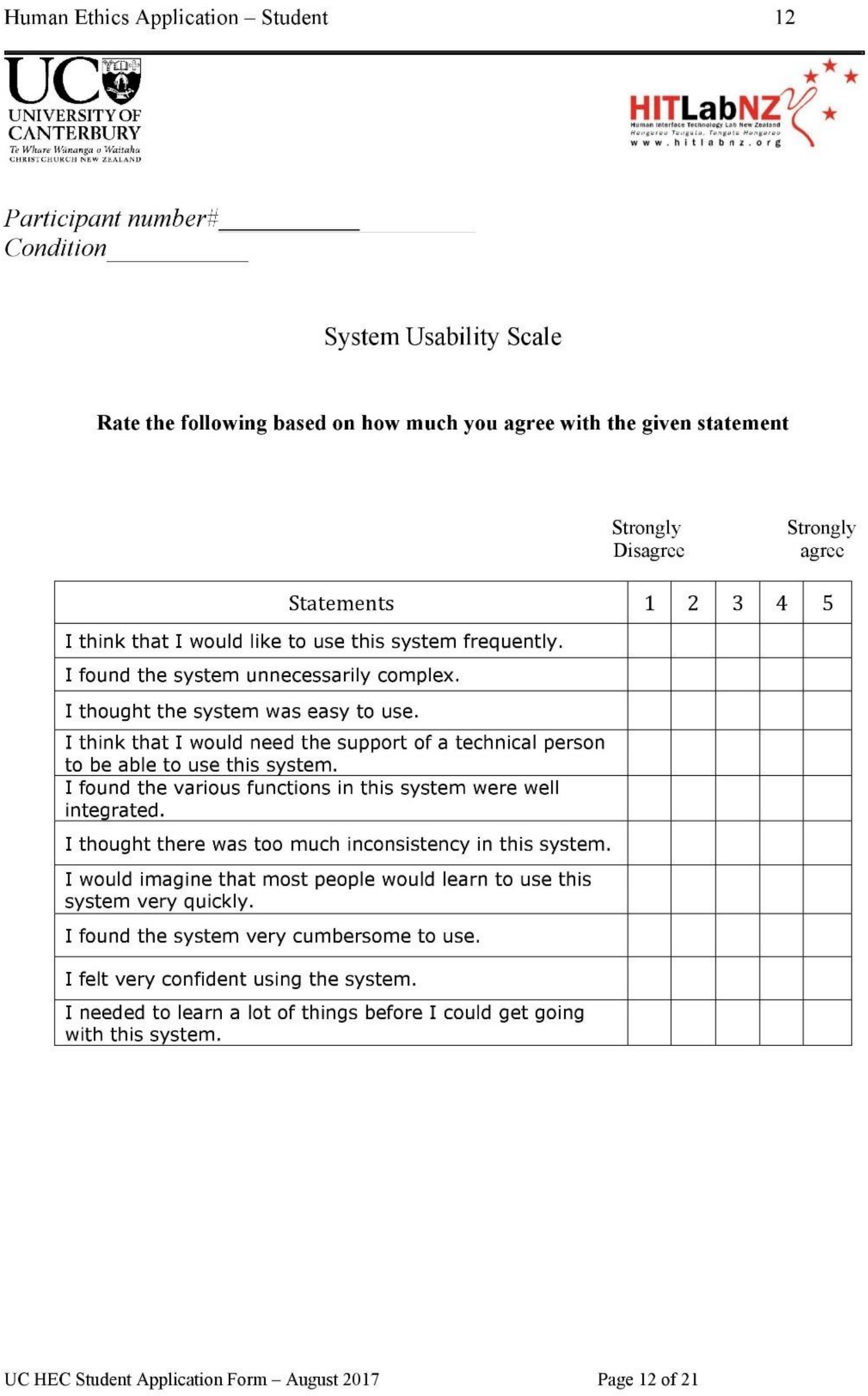


5. 3. How engagement the interaction method is?

Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Extremely

B.4 System Usability Scale



UC
UNIVERSITY OF
CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HITLabNZ
Human Interface Technology Lab New Zealand
Hangarau Tangata, Tangata Hangarau
www.hitlabnz.org

*Participant number#*___________
*Condition*___________

System Usability Scale

Rate the following based on how much you agree with the given statement

	Strongly Disagree				Strongly agree
Statements	1	2	3	4	5
I think that I would like to use this system frequently.					
I found the system unnecessarily complex.					
I thought the system was easy to use.					
I think that I would need the support of a technical person to be able to use this system.					
I found the various functions in this system were well integrated.					
I thought there was too much inconsistency in this system.					
I would imagine that most people would learn to use this system very quickly.					
I found the system very cumbersome to use.					
I felt very confident using the system.					
I needed to learn a lot of things before I could get going with this system.					

B.5 Post-Experiment Questionnaire

2020/2/4 Post-Experiment Questionnaire

Post-Experiment Questionnaire

1. Participant number

____

2. Which interaction technique do you prefer more?

Mark only one oval.

○ Conventional interaction technique
○ TMR (Tap-Move-Release) interaction technique

3. Please explain your reasoning.

____
____
____
____
____

https://docs.google.com/forms/d/1fUDKGeou7P64wAIG9I-HH5a1k4cIUQMhDgzfyzjHEtI/edit 1/1