UNIVERSITY OF CALGARY

User-defined Single-hand Microgestures

by

Edwin Chan

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE

CALGARY, ALBERTA

MAY, 2017

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "User-defined Single-hand Microgestures" submitted by Edwin Chan in partial fulfillment of the requirements of the degree of Master of Science.

Supervisor, Dr. Frank Oliver Maurer, Department of Computer Science

Dr. Mario Costa Sousa, Department of Computer Science

Dr. Lora Oehlberg, Department of Computer Science

Date

Abstract

Gestural interaction has become increasingly popular, as enabling technologies continue to transition from research to retail. The mobility of miniaturized (and invisible) technologies introduces new uses for gesture recognition. This thesis investigates single-hand microgestures (SHMGs), detailed gestures in a small interaction space. SHMGs are suitable for the mobile and discrete nature of interactions for ubiquitous computing. However, there is a lack of end-user input in the design of such gestures. We performed a user-elicitation study with 16 participants to determine their preferred gestures for a set of referents. We contribute an analysis of 1,632 gestures, the resulting gesture set, and prevalent conceptual themes amongst the elicited gestures. These themes provide a set of guidelines for gesture designers, while informing the designs of future studies. With the increase in hand-tracking and electronic devices in our surroundings, we see this as a starting point for designing gestures suitable to portable ubiquitous computing.

Acknowledgements

- To Dr. Frank Maurer – Thank you for providing countless opportunities in your ASE Lab throughout the years, and for instilling confidence in me. Lacking motivation in my undergrad, with little more than my abysmal GPA to show for, I felt both aimless and isolated. You offered me a healthy environment to grow, and encouraged me to do an internship and later, a Master's degree. Thank you for always giving me reliable guidance, yet allowing me freedom and flexibility in both work and school. I have truly been in your care.
- To Dr. Xing-Dong Yang – I never thought the study we started in my undergrad would become my thesis research. I was completely new to research, and without your enthusiasm and expertise, I would not have been able to take this research to where it is now. While I was not very proactive, you ensured that we did everything possible to prepare for every deadline. Without your help, I am doubtful we would have gotten the award at CHI, and I would not have been nearly as prepared for the presentation. Thank you for your sincerity and guidance.
- To Dr. Wolfgang Stuerzlinger – Despite not having met in person until more than a year after the research started, you always gave invaluable feedback at every stage. I still remember sending out a draft at 3AM, only to receive an edited version at 4AM with innumerable comments. Your commitment and attention to detail definitely helped to present the research more favorably, and the experience you brought to the team was invaluable to our success. Thank you for being a vital part of my few short years in research.

- To Teddy Seyed – I already promised you my first born son, so not much more to say here man. You helped me when I joined the lab, and you always had more confidence in me than I did myself. You guided me through my first paper, and you convinced me to do a Master’s. You were my go-to when I had any grievances, and you always gave me sound advice and encouragement. We had some fun times at the conferences, and I’m sure it doesn’t end here. I’m glad to have you as my friend and mentor.
- To Chris Burns – When I wanted to join the ASE Lab, I had no idea how I would do so, with nothing that would help me stand out. I remember contacting you about the research you were doing, and I ended up having a lot of fun working with the Kinect. I found some motivation for the first time in years, and I am so thankful that you and Teddy vouched for me to join the lab. You left the lab shortly after I started, but you still continued to give me so much advice, especially into industry. You also found the time to take us around during CHI, despite how busy your work must have been. Last but not least, between sleeping on the floor and sleeping with a guy I didn’t know, your air mattress was truly a light unto the darkness.
- To Yuxi Wang – Yuxibro, I still remember being disappointed after finding out the Yuxi joining our lab was not a girl. How could I have known then, that you would become one of my best bros? The countless times we worked in the lab until early morning, those were probably some of my most productive days during my graduate life. We went fishing in the game, then we went for real fish for dinner, and then we ordered the fishy fish, before that lady gave us the seared fishy fish. Man, what a journey. Let’s go for fish again soon dude.

- To Stephanie Law – You have been my close friend for a dozen years now, so I don't really need to tell you how much I appreciate you. However, I must thank you for drawing all the hand gestures for me, because without those, there would only be text; nobody wants to read that. I've had several researchers (even from overseas!) ask me which program I used to create these gestures, to which I gave the snarky reply: my friend.
- To Theodore Hellmann – Thank you for teaching SENG 301. Your enthusiasm was apparent in your lectures, and you tried to make everything clear and relatable to us students. It was refreshing, and your class was something I looked forward to. In fact, your class was the only one, in my life, where I did not miss a single minute of lecture. I was happy for you when you won the teaching award, and I like to think that my positive experience in your class motivated me to become a good TA later on. I also have to thank you for inviting us to the ASE Lab open house. Without that, none of this would have happened.
- To Robin Arseneault – We were so fortunate to have you at the ASE Lab, supporting us every way possible. You handled all the administrative paperwork, organizing events, procuring equipment, ethics approvals… the list goes on, and it was with your help that we were able to focus on school and lab work. I appreciated every little thing you helped us with, but all of it combined really simplified our lives to a tremendous extent. Also, impeccable taste in selecting the appetizers for the SurfNet workshop!
- To Mahshid Marbouti –We never worked together directly, but it was always nice to see you in the lab, and to chat about random grad stuff. You also gave me a lot of advice about TAing, and gave me a lot of reassurance. In my last semester of TAing, the slides

you provided me with were invaluable, saving me lots of headache when preparing each tutorial. Best of luck for your own studies and career, and thank you for all the help you have given me!

- To ASE Lab undergrads – Irene, Jade, Dianna, Dan, Cooper, Alec: The lab would not have been the same without all of you. Sure, the lab was LOUD at certain times, but it's much better than sitting at a desk by myself every day. It was nice to see all of you working hard on the projects, and having fun both within and outside the lab. You are all very talented people, and I am glad I was able to work with all of you. Best of luck in the future, and hopefully you will all become mentors to the new undergrads. To the new undergrads (Rachel, Austin, Teresa, Steven): It looks like all of you will fit in just fine, and carry on the lab traditions. I hope that all of you will not only contribute to the lab, but also get as much out of the lab as I did. I think it is a tremendous opportunity, and I'm not sure it's possible to find a better supervisor than Frank.
- To the students that I TAed – You have all caused me tremendous headaches. Seriously. I am convinced that all of you try to make it as difficult as possible to grade your assignments, from missing files to incorrect file names, and alas, files that do not compile. However, it brings me some degree of consolation that I was able to impact your learning, hopefully in a positive way. I hope that I have been a satisfactory mentor to all of you, and that you will all succeed in your studies and future goals. As they say, the greatest reward for a teacher is for a student to understand and succeed… or something. Every time I graded a good assignment, I secretly thanked you and was happy for you. This was especially true for the students that I know were struggling, but turned it around through your hard work and determination. Despite the many

hardships, I truly enjoyed teaching all of you. I just wish you told me sooner when I had the wrong screen on the projector. And please, no more Fifty Shades of Grey hidden in your assignment documentation.

Dedications

I dedicate my MSc work to my parents. Despite our rocky relationship at times, you have always supported me in your own ways. Having a place to call home, being relieved from the household chores, and always returning to find a meal on the counter; despite not being very vocal, I do sincerely appreciate all of the things you do for me. Thank you.

Publications from This Thesis

Portions of the materials and ideas presented in this thesis have appeared previously in the following peer reviewed publications:

- *User Elicitation on Single-hand Microgestures.* Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Yang, Frank Maurer. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI 2016). ACM. New York, NY, USA, 2016.

Table of Contents

List of Figures

Chapter One: Introduction

Throughout the history of computing, interaction methods have continued to evolve alongside technological advancements. From the age-old punch cards, to the omnipresent touchscreens, we have constantly introduced new methods to interact with our devices. However, inherent in all the countless unique interaction methods is the recurring theme of simplification. Short-cuts, aliases, hotkeys, macros; these are terms which frequently come up as ways to reduce the time and complexity of an interaction. In particular, gestural input has repeatedly been used as a form of simplification. Drag-and-drop, two-fingered scrolling, swiping away notifications, these are all gestures that have been developed to empower the user while simplifying the usage of a device. In recent years, the miniaturization of technology has made portable computing more accessible to the consumer, with “wearables” and “Internet of Things” becoming trending buzzwords. New form factors mean new interaction methods. One challenge to interacting with these smaller devices is the diminishing interaction space. As our devices continue to shrink, to the point where they are sometimes hidden and woven into our daily lives, we have to wonder: what and where can we interact with?

One solution is to transfer the interaction space from the device to the user itself. Several techniques have been proposed to detect skin-based input [53,16,38,24,46], where touching an appendage to another part of the body serves as an input modality. There is a major advantage of skin-based input over traditional methods, since there is no (a-priori) need of any apparatus acting as a medium. While an input recognition device is still required, the body itself becomes the medium being acted upon. Skin-based touch gestures such as tapping, pinching, and swiping, have been explored on several parts of the body. Most commonly, palms and forearms were used as touch surfaces given their relatively flat anatomy and perceived accessibility [24,38,56]. However, skin-based input is often restricted by clothing choices, in part due to the weather as well as cultural and religious practices. The hands are usually exposed, but always carrying an input device can be inconvenient. For example, smartwatches are worn on the wrist, while music players are often worn using an armband. Recent improvements in gestural recognition technology have made it possible to detect single-hand gestures, using devices which do not occupy the hands (eg. rings, wristbands, armbands). These single-hand gestures are promising for enhancing interactions for ubiquitous computing, and should be carefully studied and implemented.

This chapter provides an introduction of this thesis. Section 1.1 provides a brief overview of gestural input, while Section 1.2 explores the usability and user experience associated with gestures. Section 1.3 defines single-hand microgestures (SHMGs), the main topic discussed in this thesis. Section 1.4 motivates our research, while Sections 1.5 and 1.6 list our research questions and research goals respectively. Section 1.7 introduces the research methodology to be used. Section 1.8 summarizes the contributions of this thesis, and Section 1.9 describes the structure of this thesis.

1.1 Gestural Interaction

Although gesturing is an interaction method itself, countless other interaction methods have attempted to incorporate some form of gestural input. Gestures are a natural way for people to interact with not only the devices around them, but also the people and even the world surrounding them. As described by Kurtenbach and Hulteen [21],

> *"A gesture is a motion of the body that contains information. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed."*

This implies that gestures are symbolic, and inherently represent and convey some information. Gestures such as waving goodbye are commonly used as a non-verbal method to quickly communicate this inherent information. It has been shown that gesturing can help us communicate more effectively, both on its own and when used concurrently with speech [20]. Gesturing can also improve cognitive abilities, enhancing our ability to remember and recall information [48,12]. Used frequently in everyday situations, gesturing is a natural and integral method of communications. Some examples of common gestures are included in Figure 1, although depending on the culture and context, these gestures may have additional meanings.

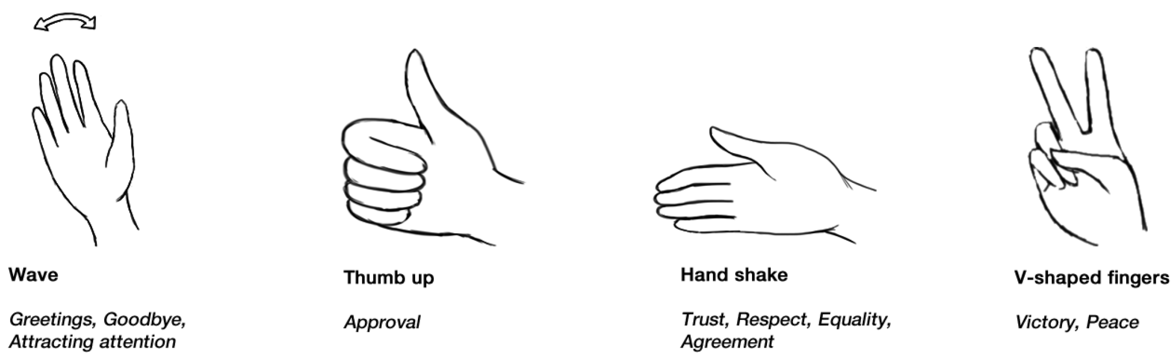

Figure 1: Gesturing is a part of our daily interactions. Some examples include a) waving the hand, b) pointing the thumb up, c) shaking hands, and d) making a V-shape with the fingers. Gestures can have often have more than multiple meanings, depending on context and culture.

Although gesturing is natural and can be used to great effect, there are some obvious drawbacks as well. Something which is easier to learn does not also imply discoverability nor memorability. For example, a keyboard shortcut may not be obvious to begin with, even though it can be easily learned and performed. Also, a single gesture can symbolize various meanings, depending on who interprets it. In different contexts, a simple waving gestures can represent both goodbye as well as declining something. In different cultures, the V-shaped "peace" symbol may represent "victory", or even have vulgar connotations. When considered alongside the often unreliable detection of gestures by computers, it becomes obvious why not all gestures are natural or effective methods of input.

Figure 2: Devices such as the Apple Watch Series 2, Fitbit Alta, and Smarty Ring have small touch controls. Interacting on these devices can be difficult due to the limited interaction space. (Images are not to scale.)

When considering small devices such as smartwatches, physical space for interactions and controls is usually limited [57,55] (Figure 2). Touchscreens tend to be very small, and very few buttons or keys can be placed on the device. In such cases, mid-air gestures are one of few alternatives available to interact with a device. In particular, a mid-air hand gesture is suitable for ubiquitous computing, since it is more discrete than large movements such as waving your arms. A mid-air hand gesture still requires an input device for recognition, but decouples the interaction space from the input device itself (Figure 3). Instead, the hand and the space surrounding it becomes the interaction medium. This makes mid-air hand gestures more suitable in portable computing which is dominated by smaller, mobile devices.

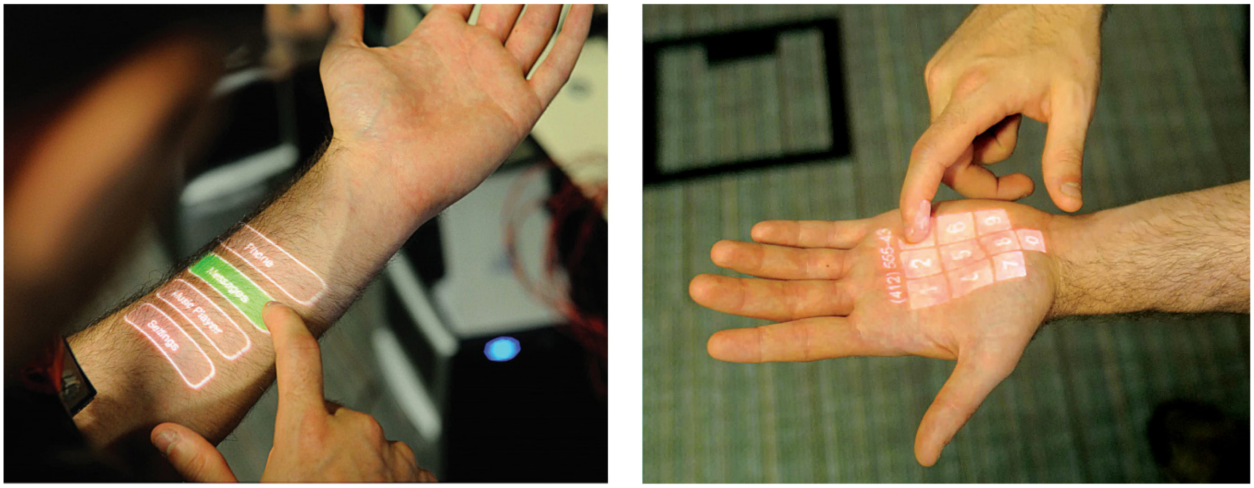

Figure 3: Skinput by Harrison *et al*. uses an armband to measure acoustic input, enabling interactions directly on the skin [16].

Given the uses as well as shortcomings of gestural input, there is both an incentive to use gestures and a need to design "good" gestures. This is often defined by the "discoverability, ease-of-performance, memorability, or reliability" of a gesture [31]. All of these are measured from the user's perspective, thus it is especially important to involve users in the design of gestural inputs.

1.2 Single-hand Microgestures (SHMGs)

While innumerable gestures exist, different types of gestures have their own usages. In the context of ubiquitous, portable computing, gestures should be always available to the user. Ideally, they should not interfere with the activities of the user, and can be performed quickly and in parallel with any ongoing activities. We define and investigate single-hand microgestures (SHMGs) as a suitable gesture type for ubiquitous and portable computing.

Figure 4: A variety of single-hand microgestures (SHMGs) are shown here. SHMGs are performed using the finger(s) to interact with the rest of the hand.

A SHMG is unique relative to other touch inputs or microgestures, resulting in its moniker. Examples of SHMGs are seen in Figure 4. Traditional touch-input is performed by a user, often with their hands, on some sort of detection device, such as a digitizer or a camera-driven touch sensor. The interaction is done directly on the device, which means a device must be retrieved or produced before an interaction can be made. In comparison, a single-hand gesture is defined here not only as performed by a single hand, but also performed on that same hand. While a device is still needed to recognize each gesture, the interaction surface is removed from the input device, keeping the hands free for other tasks. This is significant because it allows the gesture to be performed anytime and anywhere. When an interaction is not being performed, the device can continue to wait for gesture input without being held or prepared otherwise. In addition, the gesture can easily be performed secondarily with one hand while performing another task. Several studies found that users overwhelmingly preferred single-hand gestures over bimanual ones [59,46,19]; users were observed mirroring gestures on either hand to adapt to different contexts [43].

The single-hand nature of SHMGs relates to the microgestures designation as well. As interpreted by Wolf *et al.*:

> *"We understand microinteractions as interactions that are task-driven and goal oriented, and which may include system feedback. They can be evaluated with traditional usability*

> *metrics such as effectiveness, efficiency and user satisfaction. In contrast, microgestures are actual physical movements, e.g. of fingers, which are recognised by the system, and where the system reacts upon. Microgestures are part of microinteractions. Within the related work of microinteractions, the main focus is on short-time manual motor interruptions, or on manual synchronous tasks."*

In adherence to this interpretation of microgestures, SHMGs are both subtle and purposeful. Requiring only a single hand means that SHMGs can easily be performed as a secondary task, in parallel with other primary tasks. The interaction is limited to the small space of a single hand, within a short timeframe. As such, SHMGs can be performed naturally in public contexts where large or prolonged gestures may be perceived as socially awkward [41].

1.3 Motivation

When evaluating a gesture set, we want to consider usability from the end-user's perspective. We should consider whether gestures are easy to perform and remember, intuitive, metaphorically logical, and ergonomic [35]. Nielsen *et al*. compared the traditional technology-approach to designing gestures, with a human-based approach. The traditional technological approach of choosing gestures is to choose a set of gestures which are easily recognized by the system, before applying them to an application. The resulting gestures were often ergonomically stressful to perform, and lacked a logical connection between the gesture and the associated functionality. Using the human-based approach, gestures were instead elicited from users, with a focus on the usability principles listed above. After benchmarking the results from both approaches, the authors concluded that the human-based approach led to an easy-to-remember gesture vocabulary. It was also "fast for the testees to learn and remember [the gesture vocabulary]".

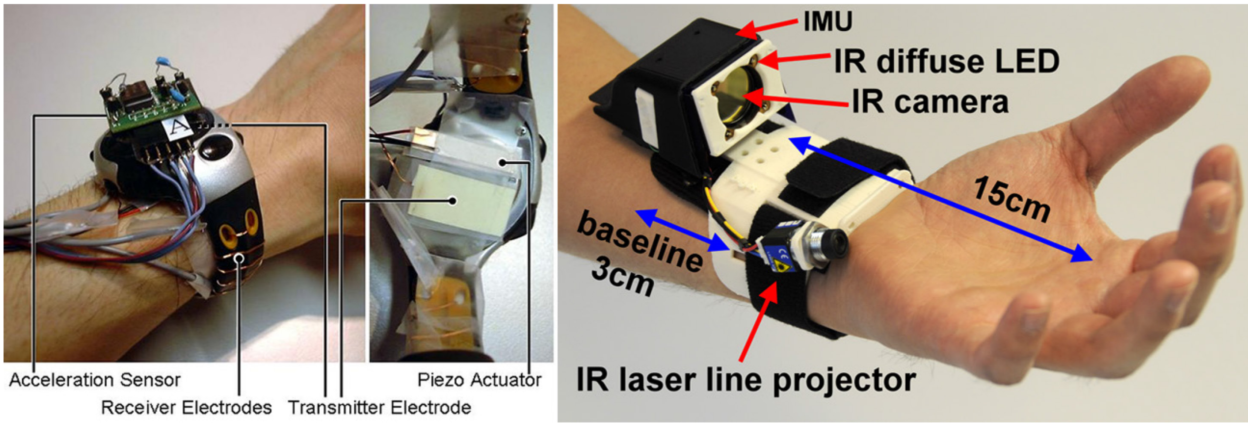

Figure 5: Existing research such as GestureWrist [40] (left) and Digits [18] (Right) has focused on gesture recognition technologies. These devices make SHMG recognition possible, but thus far SHMGs have not been elicited from users.

It is then logical to involve users in the design of SHMGs, to define a gesture set with is easy to perform and remember, intuitive, metaphorically logical, and ergonomic. However, thus far there has been no user-elicited SHMG gesture set, even though SHMGs have been discussed as both primary and secondary topics by other researchers. Most of these discussions focus on enabling technologies such as body-mounted cameras [30,26,18] or sensors [29,24,40] (Figure 5). One study elicited a gesture set from experts [60], but the gestures were performed in several contexts where users are gripping an object, such as a pen or a steering wheel. No elicitation study with end-users was performed for hands-free single-hand gestures. We aim to address this lack of user input in the design and implementation of SHMGs.

This user input will be particularly helpful to interaction designers and implementers, given the recent advancements in enabling technologies. Miniature, portable devices are now capable of accurately and reliably detecting SHMGs [3,25], making SHMGs a particularly interesting topic of study.

1.4 Research Goals

This thesis has three primary goals. The first is to inform our target audience: future designers and implementers of SHMGs. An extensive literature review will be conducted to understand the current state of SHMG recognition technology. The second goal is to determine the best method for involving users in the design of SHMGs, using a human-based approach to gesture design. Existing user studies on hand gestures will be referenced, to identify a suitable methodology for creating a user-defined gesture set of SHMGs. The third goal is to provide a useful set of guidelines for the design of SHMGs, to improve subsequent studies and implementations of SHMGs. These guidelines will be extracted from the results of the elicitation study.

1.5 Research Questions

To demonstrate the relevance of our research, and to provide a clear contextual overview of the research space to potential designers of SHMGs, we begin with this first question:

1. *What is the current state of research regarding SHMGs?*

Answering this question will achieve the first research goal, by informing the audience of the SHMG research space. Despite being discussed as primary and secondary topics in existing work, SHMGs have never been studied with a focus on user involvement. Since this will be the first study to elicit SHMG gestures from users, the following needs to be determined:

2. How can we elicit SHMGs to trigger specific tasks from users?

Although there are no elicitation studies for SHMGs, there are numerous elicitations involving other gesture types. These existing studies are valuable for both their methods and their findings. We can identify a suitable method to conduct an elicitation study of SHMGs, while comparing results to see if SHMGs exhibit properties which resemble other gestures. Furthermore, by using a similar study method, we can provide useful data for evaluating the method itself. After determining an appropriate elicitation methodology, a study can be conducted to address the third research goal. From the study results, we want to answer:

3. *What observations can be made from the elicitation study results, to help guide future designers and implementers of SHMGs?*

The results will be both quantitatively and qualitatively analyzed. Gestures will be counted to determine the frequency of each gesture for each task, and within the complete set of proposed gestures. A gesture set will also be proposed, detailing the most popular gestures for each task. Themes and patterns will be determined within the proposed gestures. User comments and existing research will be referenced to further explain any findings.

Research questions #1 and #2 are addressed in Chapter Two, through an extensive literature review. Question #3 is explored through Chapters Three, Four, and Five, which will discuss the results of the elicitation study along with any findings and limitations.

1.6 Research Methodology

We employ Wobbrock *et al.'s* elicitation methodology [28], which has been used by more than 20 studies to explore user preferences towards gestural input [52,63,10]. To better realize the potential of SHMGs, we are interested to fully understand human preferences without being restricted by technical compromises. Compared to existing studies of other gesture types, we can ignore technological limitations to a greater degree, because there is little implementation required to conduct a study. This is possible with SHMGs, because the input space or medium is separated from the recognition device. With other gestures types which rely on interacting directly on or with a device, there is a need to provide a device for the user to interact with in a study, even if it is mocked up. Since SHMGs uses the hand as its interaction medium, a user can design SHMGs exactly as they would be implemented, without regards to which recognition technology is used. For example, a recognition device using either a camera [3] or a strain-gauge sensor [25] can be used, without affecting how a gesture is performed. The elicitation methodology is further discussed in Chapter Two, and our application of this methodology is discussed in Chapter Three.

1.7 Thesis Contributions

We provide a comprehensive summary of the research space surrounding SHMGs and gesture elicitation studies. From our study, our contributions begin with the classification of 1,680 elicited gestures (see Appendix B for the complete data set), followed by the statistical analysis of the data using Vatavu *et al.'s* revised agreement rate [25]. We conclude with a set of design guidelines that offer qualitative insight into end-user thinking when designing SHMGs. The versatility of SHMGs make them suitable to many scenarios, and we see our work as a preliminary effort to designing better SHMGs for enhanced adoption and user experience.

1.8 Thesis Structure

This chapter provided a concise background about the research for this thesis. The primary research topic, SHMGs, was defined in comparison to other gesture types. The research goals, questions, methodology, and contributions were outlined.

Chapter Two: Background and Related Work – provides an overview of related research. The first part catalogs existing research on SHMGs, while the second part reviews existing studies and methodologies for conducting gesture elicitation studies.

Chapter Three: User Elicitation on Single-hand Microgestures – documents the design of the elicitation study, the analysis techniques, and the quantitative results.

Chapter Four: User Comments and Design Implications – combines the feedback obtained through the interviews at the end of each elicitation, with the quantitative analysis. Themes and motifs identified from the results help form the design guidelines for SHMGs.

Chapter Five: Study Limitations – lists the challenges and restrictions encountered by our elicitation study.

Chapter Six: Conclusion and Future Work – discusses the importance of our contributions, and introduces future work which can be extended from our research. The research questions and goals are revisited, and a conclusion is made regarding the research of this thesis.

Chapter Two: Background and Related Work

2.1 SHMG Technology

The decision to study single-hand microgestures (SHMGs) resulted from a series of observations, when looking at the current state of gestural computing. From a commercial perspective, 3D or mid-air gestures only started becoming more prominent in the last decade, with the introduction of low-cost gesture recognition devices such as the Microsoft Kinect (2010) [27] (Figure 6). Before that, users almost mostly relied on mouse and keyboard input, with touch input used on many mobile devices such as phones and tablets.

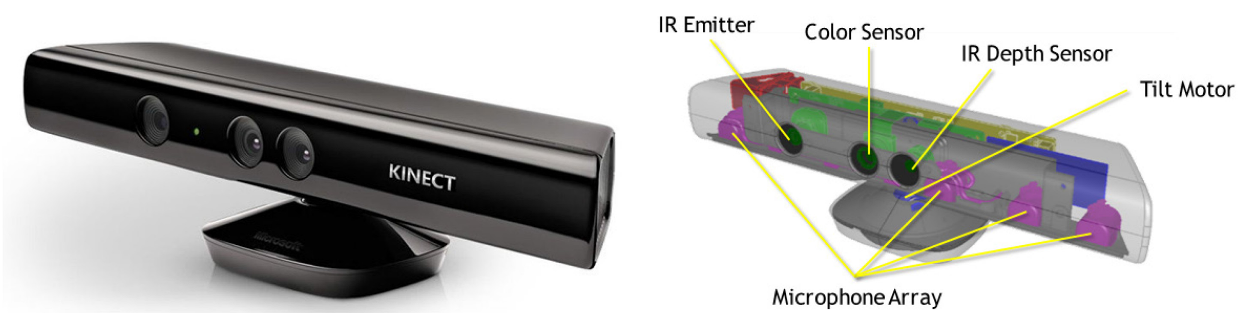

Figure 6: The Microsoft Kinect, released in 2010, was a relatively low cost device combining both RGB and IR cameras. It significantly lowered the cost of implementing visual gesture recognition.

More recently, other commercial products such as the Leap Motion [23] (2013) and Myo armband (2015) [51] (Figure 7) have started to appear, offering more options for detecting SHMGs and other hand gestures. Although we have yet to reach a point where hand-gesture input is ubiquitous and widely accepted, there has been an increase in hand gesture recognition devices nonetheless.

Figure 7: The Myo armband (left) detects hand gestures through electromyographic (EMG) sensors [51], while the Leap Motion (right) relies on infrared stereo cameras [23].

As motivated previously, single-hand gestures are preferred by users over bimanual gestures [19,46,59]. Furthermore, a social acceptance study on gestures by Rico *et al*. emphasized the need for gestures to be discrete [42]. "Gestures that required the participant to perform large or noticeable actions were the most commonly disliked gestures." Although SHMGs by definition are both single-hand and discrete, and should be preferred by users, they have not been specifically researched nor widely adopted. This was mostly due to the lack of enabling technologies, thus it is important to first look at the current state of SHMG research.

2.1.1 Portable and Ubiquitous Computing

Even in the early days of computing, there has always been a desire to make computing portable and ubiquitous. Technological advances have allowed computers to shrink, from room-sized behemoths to desktop computers, to mobile devices such as laptops, smartphones, and tablets. With the advent of the Internet-of-Things movement, tiny computers are commonly embedded into the most mundane of everyday objects, such as smart glasses, smart cars, smart trashcans, and even smart water bottles. There seems to be a desire to make computers a part of our daily lives, by making them invisible yet omnipresent. With devices constantly growing smaller, an increasingly apparent problem is the lack of real estate for interaction controls, such as buttons or touch screens [55,57]. This problem has motivated various types of research, including skin-based interaction techniques [38,24,56], and mid-air gesture recognition [53,14,18]. In both cases, the recognition device is no longer directly interacted with; instead, the user's body becomes the interaction space itself.

2.1.2 Skin-based Inputs

On the surface, using skin-based inputs appears to be a very logical answer to interacting with small devices. Interactions are done directly on the body itself, and the interaction space is naturally always "with" the user. Skin-based interaction research often studied touch inputs on the forearm, given its accessible location and flat anatomy (Figure 8). Skinput appropriated the forearm as a touch surface by measuring acoustic energy traveling through the forearm [16]. Takemura *et al.* used a similar approach, but focused on sound transferred through the bones [50]. The SonarWatch combined an ultrasonic rangefinder with a capacitive touch sensor in the form of a wristwatch, to enable touch gestures on the forearm [24]. Palm+Act used an RGB camera to estimate the force of a touch input, but used the palm instead as the interaction space [38]. Many other skin-based approaches exist, including OmniTouch [15], SenSkin [36], and AugmentedForearm [37].

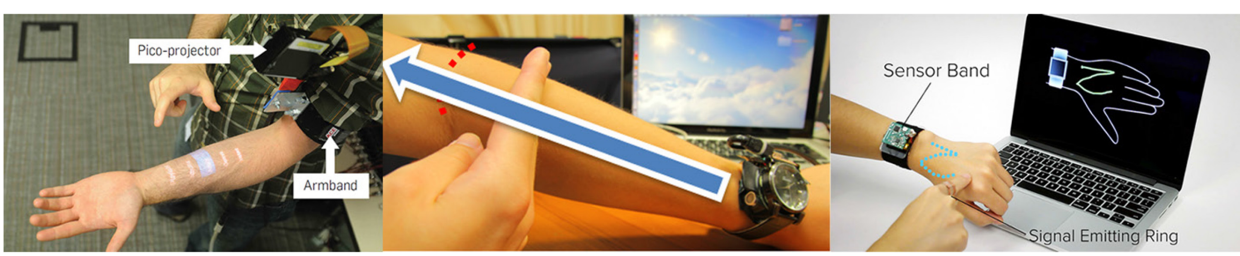

Figure 8: Skinput (left) [16], SonarWatch (centre) [24], and SkinTrack (right) [65] use arm or wrist-worn devices to recognize touch interactions performed directly on the skin.

In 2014, an elicitation study was conducted to understand user preferences for skin-based inputs [56]. All of these examples focused on touch input, and used the human body to extend the interaction space of a device. However, despite growing interest in skin-based input techniques, a commonly cited limitation of many such methods is the need for the skin to be exposed. For example, wearing a jacket on a cold day would cover the forearm, making these techniques impractical for daily use [37]. Compared to the forearm, the hand is less likely to be obscured, and is more suitable for a wide range of contexts.

2.1.3 Early hand-gesture detection

The use of hand gestures in computing is not a new idea by any means [22]. Many implementations of gesture recognition exist, and another work lists at least 37 such implementations just for 3D gesture recognition [9]. Researchers continue to present improved technologies and new implementations, and implicitly make gestures easier to adopt and use by users. As far back as 1977, there have been attempts to track hand movements as input commands using a data glove [8]. Despite inadequate accuracy and precision to recognize complete gestures, the device could still be used to manipulate 2D widgets such as sliders.

Figure 9: Data gloves such as the VPL DataGlove and Mattel Power Glove have been used since 1977 as hand gesture input devices.

One of the earliest devices capable of recognizing hand gestures, including some SHMGs, is the Digital Data Entry Glove designed by Gary Grimes in 1981 [13]. Using touch and proximity sensors, it could determine if the user’s thumb was touching another part of the hand or fingers.

This device was created specifically for alphabet input, and was not adapted for generic gesture recognition. One of the most-discussed data gloves in literature is the DataGlove developed by VPL Research, a general-purpose interface device offering 10 degrees of freedom with the ability to track simple hand gestures [66] (Figure 9). In 1989, the Power Glove was developed by Mattel for Nintendo as arguably the most well-known data glove [49]. Although it had low accuracy and could only track simple gestures, the Power Glove was relatively inexpensive and was marketed publicly to consumers.

Even with high enthusiasm for researching and developing data gloves, there were still many obvious drawbacks preventing them from being widely adopted for general use. At the time of a 1999 survey, the relatively low-end 5DT Data Glove ™ started at $1030/pair. On the opposite end of the spectrum, the SuperGlove sold for $5,000/pair, with a wireless option for $20,000 [22]. Besides cost, there were still many other factors which limited the adoption of data gloves. Like any regular glove, data gloves can restrict hand movement. Hand size and anatomy are also important factors which lead to lower accuracy, requiring devices to be calibrated. Due to the obscurity of mobile computers at the time, data gloves were essentially tethered for stationary use. Finally, existing technological limitations meant the gloves were rather large and cumbersome.

2.1.4 Vision-based Gesture Recognition

An alternative to glove-based solutions is vision-based gesture recognition [62]. Using either dedicated cameras or built-in ones such as webcams, a user's hand motions are captured [40,18]. The video is then processed through computer vision and learning, and classified into unique gestures. Computer vision offered several advantages over data gloves. First, general-purpose cameras were much cheaper when compared to the data gloves at the time [22]. Cameras could be also be used for taking photos or video recording, while data gloves only served as gesture input devices. Second, the user is not required to wear any device, resulting in higher mobility and a more natural feel. However, vision-based systems were not perfect either. Albeit less cumbersome than data gloves, the input space using cameras is still limited to the view and range of the camera device. The placement of the cameras has a severe impact on the detection accuracy. In the worst case, cameras can be completely obstructed, meaning users have to be extremely aware of their position and orientation relative to the cameras [45].

Figure 10: Visual gesture recognition devices such as the (from left to right) Kinect for Windows v2 [27], Vicon MX-F40 [54], Creative Senz3D [6], and the Leap Motion [23] often use infrared cameras. (Images are not to scale.)

Despite their drawbacks, vision-based solutions were still highly popular (Figure 10). Solutions could cost as much as hundreds of thousands of dollars [54], or as little as $80USD [23]. Many of these solutions were also consumer-ready, making them easily accessible. In 2010, the first version of the Microsoft Kinect [27] was released for use with the Xbox 360 [28]. Soon afterwards, the device was hacked by the community to work with PCs. This Kinect camera was a $150 infrared and RGB camera, capable of measuring depth and seeing the environment in 3D. The device quickly gained prominence in the research communities, resulting in the subsequent surge of papers focusing on mid-air gestures from 2011 onwards [14]. In a 2016 systematic literature review of 3D mid-air gestures, 65 publications were analyzed, with the vast majority utilizing some camera-based system (and especially the Kinect) [14]. Of the 65 papers, only five were published prior to 2011. Despite the rapid growth and adoption of vision-based techniques, the problems restricting vision-based solutions still remained. The view range and angle of cameras were limited, and poor placement or obstruction of the cameras prevented vision-based methods from being used in many contexts. Body-mounted camera solutions were then proposed to mitigate the above problems [26,29]. Placing a camera on the body meant that the camera would be much closer to the hand, and could move with the body. The camera was less likely to be obstructed by external objects, while the hands are more likely to be within the view of the camera. At such close distance, even low-resolution cameras could produce comparable results to a more expensive camera placed far away. Earlier examples of this approach include mounting a camera on a baseball cap or as a pendant [47], mounting a Kinect camera on the shoulder [15], and mounting a CamCube 2.0 camera on the chest [26]. However, the sheer size of the cameras used made them impractical.

2.1.5 Enabling SHMGs

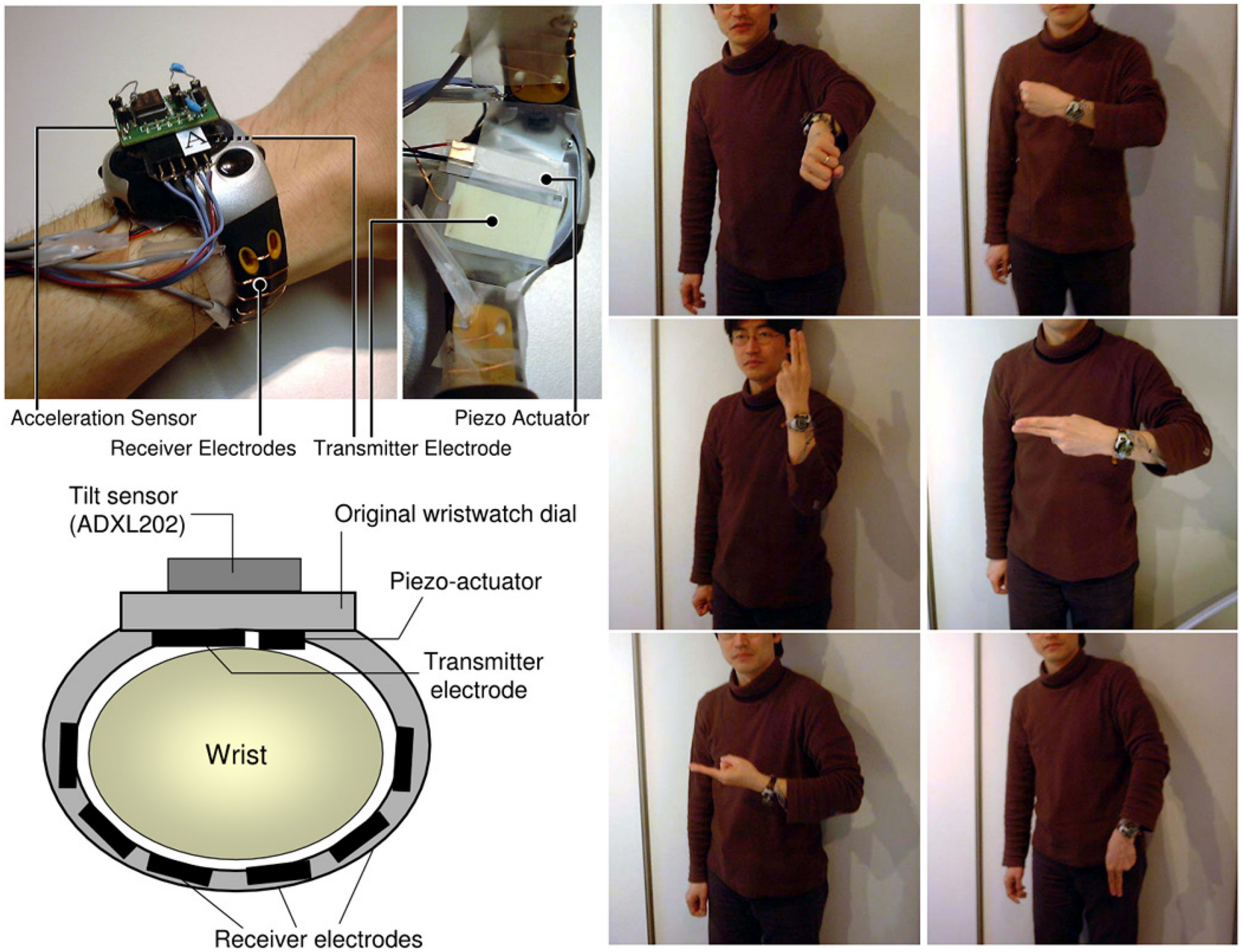

Figure 11: Rekimoto is performing hand gestures (right), using his GestureWrist device (left) [40] The device relies on a capacitive sensor, paired with an acceleration sensor.

Fortunately, with shrinking component sizes leading to the proliferation of wearable devices (or wearables), many of the problems found in both data gloves and vision-based techniques can be reduced or resolved. In 2001, Rekimoto *et al.* proposed GestureWrist, a wrist-worn device which was capable of recognizing certain hand gestures through acceleration and capacitance sensors [40] (Figure 11). Despite still being larger than a traditional wristwatch, GestureWrist was notable because it was one of the earliest solutions which relied on a wrist-mounted device, rather than a glove-based solution. In 2009, Saponas *et al.* presented an EMG approach, which enabled hand and finger gesture interaction [44]. This was a departure from existing research which "primarily focused either on using a single large muscle (rather than the fingers)…and/or on situations where the hand and arm are constrained to a surface.”

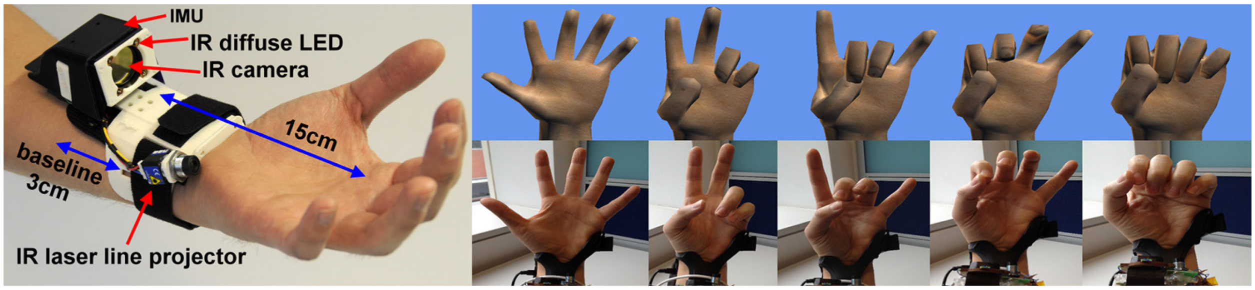

Figure 12: Kim *et al*.'s Digits device uses a combination of wrist-mounted sensors to detect hand gestures (left). It was tested with a variety of gestures and hand poses (right) [18].

In 2012, Kim *et al*. proposed a wrist-worn IR camera system, capable of detecting the 3D pose of the hand [18] (Figure 12). In 2015, the Myo armband was released commercially, using EMG sensors in combination with an inertial measurement unit (IMU) to detect hand gestures [51]. The device was small enough to wear under clothing, and connected wirelessly via Bluetooth. However, the Myo armband was released with only 5 gestures, 2 of which were SHMGs. The raw EMG data was subsequently made available, allowing developers to implement custom gestures. In 2015, two particular prototypes showed that SHMG detection can now be done both accurately and unobtrusively.

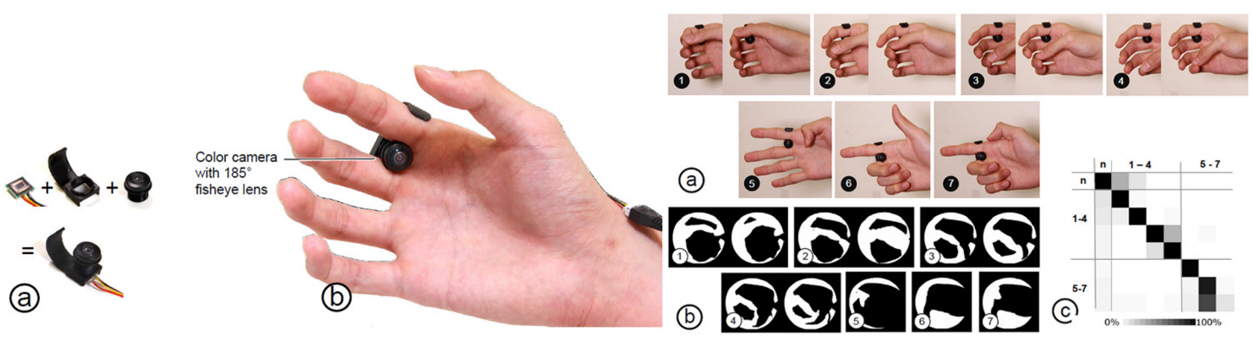

Figure 13: The CyclopsRing device uses a color camera with a fisheye lens to detect hand gestures (left). It was tested with a variety of hand gestures, many of which can be classified as SHMGs (right) [3].

The CyclopsRing prototype device utilized a Pi NoIR camera module placed between the fingers [3] (Figure 13). It was able to achieve 84.75% accuracy during a study involving 7 gestures, all of which fall within the SHMG definition. On the other hand, the BackHand prototype utilized strain gauge sensors affixed to the back of the hands, and reached a 95.8% average accuracy for 16 popular hand gestures when personalized for each participant [25] (Figure 14). Although only American Sign Language gestures were tested with the prototype, the demonstrated gestures show that SHMGs can be recognized using the device.

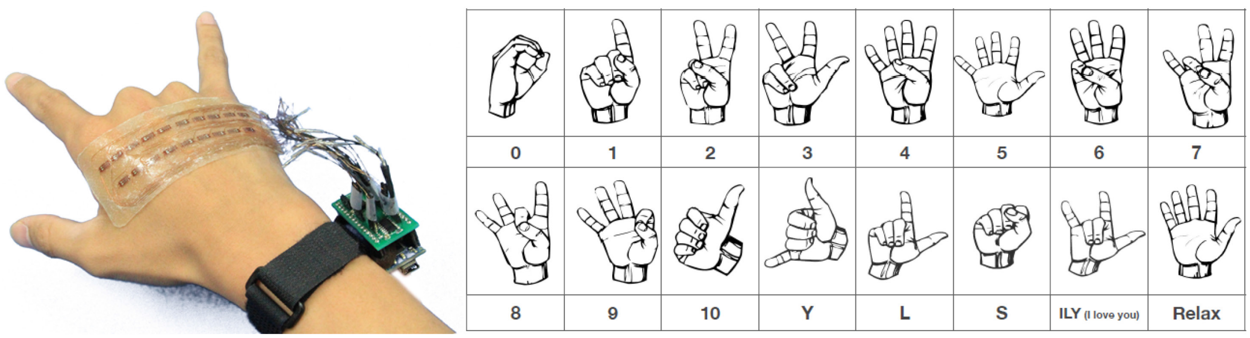

Figure 14: The BackHand device uses an array of stress gauge sensors to detect hand gestures (left). A variety of gestures from American Sign Language (ASL) and Asian culture were tested with the device (right) [25].

These prototypes have shown that SHMG detection is not only possible, but can be accomplished using small wearable devices at a low cost. However, research thus far has focused on the capabilities of the technologies, and the gestures afforded by them. User studies were mostly conducted to test the accuracy of devices by performing gestures predetermined by the authors. No prior research on SHMGs has looked at eliciting gestures and examining user preferences, despite findings which suggest user involvement leads to gestures which are more preferred [35]. A recent work proposed a taxonomy of microinteractions [60], defining microgestures based on ergonomic and scenario-dependent requirements. While the premise of investigating gestures performed in relation to hand grips (due to holding objects or devices) differed from our device-free gestures, their consultation with four experts of hand anatomy helped to define the physical traits that limit SHMGs. One participant was a sports therapist, while the remaining three were physiotherapists. Although end-users were not involved, their observations can complement user-defined guidelines and help to explain the preferences of users.

2.2 User-defined Gestures

A main concern with designing gestures is how well users resonate with such gesture sets, and whether a set of gestures is "easy to use." Unfortunately, gesture interfaces are often designed without fully consulting end-users, or sacrifice usability for ease of implementation and practical reasons [35]. As motivated by existing user studies, designers and developers often do not share the same conceptual models as the end-users that should be catered to [35]. In many cases, end-users blend concepts from other systems with which they have previous experiences [1]. These may include common household objects, phones, handheld controllers, etc. When comparing user-elicited and expert-elicited gesture sets, Wobbrock *et al.* discovered a user preference for user-elicited gestures; gestures proposed by both users and experts were most preferred by users [32]. In addition, Nacenta *et al.* found that user-defined gestures are easier to remember [33]. There is an incentivized need to involve users throughout the design and implementation of any gesture interface, which can be accomplished using various methodologies.

2.2.1 Elicitation Studies

In 2016, a systematic literature review was conducted on 3D mid-air gestures [14]. This research was particularly relevant, because it "excluded papers which evaluated mid-air gestures with the aid of devices (e.g. pen or wand) or papers that only concentrated on full body gestures (e.g. feet, legs, and torso)." This criteria is very similar to our own for SHMG, which focuses on single-hand gestures, without the aid of an external device. Although a device is still needed to recognize the gestures, the gestures themselves are performed using a single hand, and performed on the same hand. The review included 72 studies from 65 papers, using various types of *usability* and *guessability* studies. Several papers had multiple studies, while five papers did not include any evaluation. Because many papers discuss some novel implementation or device, 51 of the 65 papers presented *usability* studies, in the form of user studies (36 papers), pilot studies during design (9 studies), and field studies (6 studies). In comparison, much fewer papers conducted *guessability* studies, with 9 elicitation studies and 8 Wizard of Oz used in participatory design, where users are involved in the design process. In fact, elicitation studies can incorporate the Wizard of Oz technique [46]. In a Wizard of Oz study, the user is presented with some interface, which they interact with [7]. Since the user is unaware of the underlying implementation, this type of study can also be used to test features that have yet to be implemented, by manually showing the expected result to the user. In an elicitation study, the user is usually provided with a list of referents, and asked to perform some action on the interface/device being tested which they believe would lead to the corresponding referent being executed [35]. Many elicitation studies opt to use the Wizard of Oz technique to display the referent, when there are already some underlying assumptions regarding how the system/device should function or behave. However, an elicitation study can be conducted without the Wizard of Oz technique, if users are to be consulted before any design (or implementation) begins.

Currently, research prototypes are able to detect many SHMGs, with a relatively high accuracy [3,25]. However, no specific SHMG gesture sets have been researched nor proposed, therefore the goal should be to improve the design of all future implementations of SHMGs. The elicitation methodology, which is suited for early user involvement and does not require any working implementation, can be used to elicit SHMGs. The Wizard of Oz technique should not be used here, as motivated by a prior gesture elicitation study: "In developing such a user-defined gesture set, we did not want subjects to focus on gesture recognition issues or current motion capture technology.” [63] This was done to remove the so-called gulf of execution, which represents the discrepancy between a user’s goals and knowledge, and the level of description provided by the systems they are using [17].

While elicitation studies are used in various research domains, Nielsen *et al.* proposed a procedure for eliciting and developing user-defined gestures [35]. User elicitations using this procedure have offered contributions towards the design of the studied gestures and the overall user design process, despite eliciting unique types of gestures [10,46]. Wobbrock *et al.* discussed several intriguing concepts including dichotomous references, reversible gestures, and simplified mental models [59]. Seyed *et al.* noted the importance of aliasing gestures as a solution to varied user preferences, while offering atomic gestures and themes to help map gestures to users’ conceptual models [46]. Angelini *et al.* looked at gestures performed on a steering wheel, and observed user preferences for body parts used in gesticulation, and the frequency of gesture actions such as swipe or tap [1]. These observations are mostly of a qualitative nature however, and can be hard to measure or define.

2.2.2 Agreement Rate

To help formalize and quantify results, Wobbrock *et al.* provided an agreement measure to analyze and interpret elicited data [58]. The level of agreement is proportional to the number of participants proposing the same gestures, and the number of total participants. If many participants propose the same gestures, then the agreement will be higher. Conversely, if the proposal are extremely diverse, then the agreement rate will be much lower. If there is significant agreement between users, it would be worthwhile to determine a *consensus gesture set*. This set of gestures is comprised of the most frequently proposed gesture for each referent.

The proposed measurement method has been widely adopted by prior elicitation studies [1,30,46,53,59,60]. Since its introduction in 2009, the agreement rate has been calculated in at least 20 studies [52,4,10].

Despite being widely adopted, this formula had some shortcomings as well. Most importantly, it did not accurately represent gestures with no agreement. Unique gestures that had zero agreement trivially agreed with themselves. Therefore, gestures with zero agreement actually did not have an agreement rate of 0. The original formula also did not account for the degrees of freedom; a gesture with 15/20 matching entries had the same agreement rate as a gesture with 30/40 matching entries, despite the latter clearly showing greater agreement for the consensus gesture [52].

$$A(r) = \sum_{P_i \subseteq P} \left(\frac{|P_i|}{|P|} \right)^2$$

In the original agreement formula *A(r), P* denotes the set of all proposals for referent *r*, |*P*| is the size of the set, and P_i comprises the subsets of identical proposals from *P*. Vatavu and Wobbrock later addressed these issues with a revised formula [52], which gave a better representation of the elicited gestures. This revised agreement formula is given by *AR(r)*:

$$AR(r) = \frac{\sum_{P_i \subseteq P} \frac{1}{2} |P_i|(|P_i| - 1)}{\frac{1}{2} |P|(|P| - 1)}$$

In the same publication, the new disagreement rate and coagreement rate were introduced to quantify the relationship between multiple referents, based on the respective gestures proposed by users. The disagreement can help to dispute existing assumptions of similarities between certain referents. The coagreement rate can conversely help group seemingly unrelated gestures, and identify the shared properties between them. The coagreement rate has been used to quantitatively assert the relationship of dichotomous pairs, opposing referents which tend to have related gestures. For example, *volume up* may result from a *swipe up* gesture, while *volume down* may result from swiping down in the opposing direction. Considering the popularity of the original agreement measurement, and the advantages of the revised formula, the agreement rate measurement will also be useful for analyzing the results of an SHMG elicitation study.

2.2.3 Legacy Bias

While elicitation studies are ideal for involving users in early stages of design, no study method is without its shortfalls. An elicitation study heavily emphasizes the preferences of end-users, rather than the assumptions of authors and designers. However, each user has their own unique experiences, which directly contribute to their preference of interfaces, including gestures. Known as *legacy bias*, this problem is often encountered when designing new user interfaces. When faced with something new and unfamiliar, users instinctively fall back on prior experiences which are both familiar and reliable. Similarly, when asked to design new gestures during an elicitation study, users often relied on their past interactions with WIMP (windows, icons, menus, and pointing) interfaces. This means that the elicited gestures may not be maximizing the potential of the new interface, while users that do not share the same experiences may not be able to adapt to the proposed gestures of other users. In 2014, Morris *et al*. elaborated on the legacy bias dilemma, while proposing three techniques for elicitation studies which may reduce the effects of legacy bias. The three techniques are *production*, *priming*, and *partners.*

Production entails eliciting the user for multiple gestures, for each referent. Users then select one of these as their preferred gesture. Even if a proposed gesture is heavily influenced by the user's unique experiences, the other choices are likely to be less specific. This can be compared to a compromise when making group decisions, such as deciding on a restaurant for dinner. If each member of the group has one specific preference, it will be hard to reach a consensus. However, when each group member offers up several choices, it becomes much easier to find a common point. Although production is mainly for reducing legacy bias, it serves more than simply being a compromise. It is often said the first option may not be the best option, and many design processes include proposing several ideas, before narrowing them down and refining them. Along the same lines, the first proposed gesture may not be the preferred one, as observed in other elicitation studies [31,46]. Finally, by forcing users to propose several gestures, users are often forced to think outside the box and further explore the capabilities of the new interface.

Priming is done at the beginning of the elicitation study, by explaining the capabilities of the new interface to the user. This is often done by having participants watch a brief demonstration of the system, to get a better feel for it. Through better understanding, users are expected to be more comfortable with the interface, and they may be made aware of some functionality they were unaware of. Users can freely ask questions about the interface before the elicitation process. Authors may also benefit from these questions, which may reflect users' expectations for the new interface.

Partners encourages multiple participants to collaborate, so that they brainstorm and propose gestures together. The main issue with legacy bias is that any user's experience may be too specific to themselves, and not be common amongst the targeted user group. When multiple participants discuss their gestures together, they will often be asked to explain why they proposed a certain gesture. A proposal is likely to be rejected by other participants if they cannot comprehend the reasoning behind the gesture. Although it is still possible for several participants to share relatively unique experiences, the probability is much lower.

By using these methods to reduce legacy bias, the resulting gesture set is expected to achieve greater consensus. The gesture set is also more likely to contain "interactions that take full advantage of the possibilities and requirements of emerging application domains, form factors, and sensing capabilities [31]." Despite introducing the *production*, *priming*, and *partners* techniques, Morris *et al.* are careful not to completely dismiss the potential of legacy bias. If gestures are being elicited for a walk-up system such as a public display, legacy bias can help to improve discoverability of gestures.

In 2015, a related study by Köpsel *et al*. proposed an approach where legacy bias is leveraged rather than simply reduced [19]. Since users often rely on prior experiences which they are familiar with, a new interface may be easier to adopt if it also supports these older interactions. In a replication of the original study conducted by Morris *et al.*, 18 users were elicited for a number of gestures. The referents/tasks involved *backward* and *scroll-down* on a webpage of a browser, *close* the browser, and the same gestures for controlling an e-book reader. Most of the proposed gestures were "wipe" gestures, using some sort of swiping motion. The users were then given the chance to change their answers, yet only three users opted to, with the remainder persisting with their original choices. Every user had prior experience with these gestures on smartphone touchscreens, so it can be said they were influenced by legacy bias. Accordingly, legacy bias can also contribute to the usability and adoption of a new interface, if enough users share a common experience. Köpsel *et al*. thus concluded that new interfaces should support both new and old interactions when possible. This can achieved by *aliasing* multiple gestures for a single referent, a technique which dramatically improves input guessability [11,58,46].

2.3 Gesture Classification

A gesture elicitation study results in numerous proposed gestures being recorded, but each gesture has to be described and classified in a consistent manner for them to be compared. The *Descriptive Labeling* technique by Nielsen has been used in gesture elicitation studies [46], wherein gestures are described by "their movement and action – postures, hand/finger positions, hand trajectories, and posture – and not be what a gesture communicates or its purpose (semantic meaning). *Descriptive Labeling* can be used in conjunction with the *Chunking and Phrasing* technique by Buxton [2], which identifies *atomic* and *compound* gestures by delimiting parts of a gesture by periods of tension and relaxation into "phrases". An atomic gesture is made up of a single phrase, whereas a compound gesture constitutes multiple phrases. *Chunking* and *Phrasing* has also been used in gesture elicitation studies, as an effective way to identify the intrinsic actions performed in each gesture [46,1,59]. The combination of Nielsen's *Descriptive Labeling* and Buxton's *Chunking* and *Phrasing* makes it easy to describe, classify, and analyze the results of an elicitation study (Figure 16).

TAP on *INDEX* with *THUMB*, then
SWIPE on *INDEX* with *THUMB*

Figure 15: We follow Nielsen's *Descriptive Labeling* method to describe each gesture, based on the movements and actions rather than their semantic meanings [35]. *Tap* and *Swipe* are too atomic actions, delineated by verbs. Combining the two creates one whole compound gesture [2].

Chapter Three: User Elicitation on Single-hand Microgestures[1]

An elicitation study was conducted to identify user preferences for SHMGs. This chapter includes the study design, the classification of the elicited gestures, and various quantitative methods of analyses including the agreement rate [52].

3.1 Study Design

3.1.1 Participants

Sixteen paid volunteers participated in the study (7 male, 9 female). Participants were recruited using email lists and word of mouth. The participants ranged in age from 16 to 39 years (Mean = 22, SD = 4.97), and came from differing backgrounds including marketing, arts, psychology, and high school students. Of the 16 participants, 4 reported having experience with microgesture devices such as the Myo armband [51] or Leap Motion sensor [23]. All participants had some experience with touch gestures, along with frequent use of devices such as smartphones, computers, or gaming consoles. Participants were not screened for finger mobility or motor skills, such as from playing the piano or learning American Sign Language.

3.1.2 Apparatus

Since SHMGs are performed by a hand and on the same hand, users did not interact with any input device. Before starting, participants were shown sample SHMGs on a laptop computer to illustrate some of the capabilities and limitations of SHMGs. For the elicitation, referents were listed on a printout, with each referent being demonstrated on the laptop computer.

[1] The contents of this chapter are based upon [4].

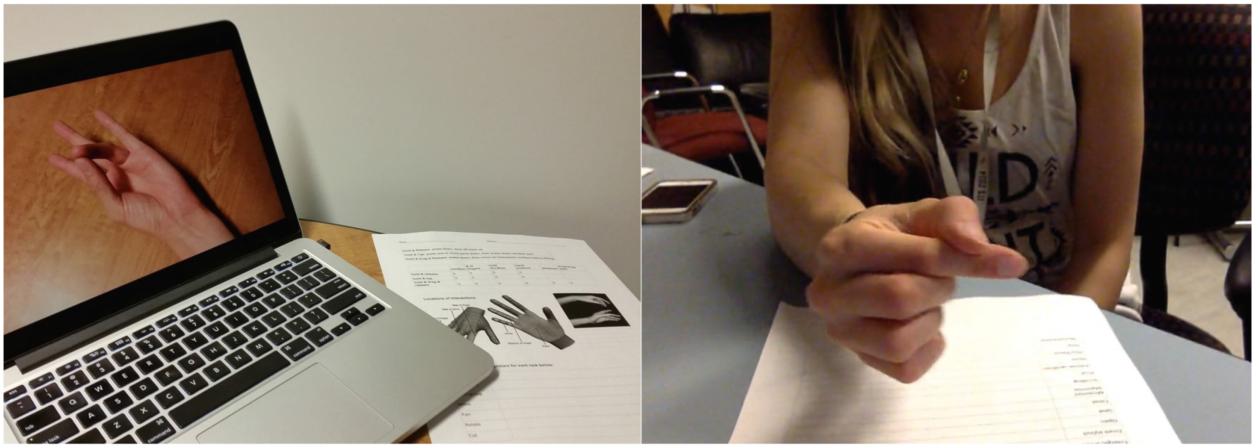

Figure 16: Diagrams of the hand and sample videos of SHMGs were used to *prime* users about the capabilities and limitations of SHMGs (left). Users were then given a list of referents (right), and gestures were elicited for each referent.

Video recording was done with a 1080p webcam mounted on a tripod, and users were able to see the live recording so they could keep their hands in view of the camera. The recordings captured each of the user's gestures, as well as communication with the users (Figure 17).

3.1.3 Referents

We wanted to form a list of common tasks which users could relate to and which they may perform frequently. To do so, we looked at Wobbrock *et al.'s* list of referents [59] as a starting point. We then added 12 tasks that are commonly performed on devices such as phones or computers. Six of these made up the Simulation category, which is commonly included in referent lists used for mid-air gesture elicitations [14]. The remaining six gestures we added are "Scroll", "Copy", "Save", "On", "Off", and "Find". In particular, "On" and "Off" were particularly important for SHMGs, because SHMGs are suitable for ubiquitous computing. Because the gesture recognition can be always available, there might be a need for users to turn the device off when necessary. The gestures were grouped into the six categories used in Piumsomboon *et al.*'s gesture elicitation for augmented reality [39]. Figure 18 lists the final set of referents that we used in the study.

Category	Tasks	Category	Tasks
Transform	1. Move	Editing	19. Cut
	2. Rotate		20. Copy
	3. Enlarge		21. Paste
	4. Shrink		22. Delete
	5. Minimize		23. Accept
	6. Maximize		24. Reject
Simulation	7. Volume up		25. Undo
	8. Volume down		26. Save
	9. Mute	Menu	27. Help
	10. Play		28. Open menu
	11. Pause		29. Close Menu
	12. Stop		30. On
Browsing	13. Pan		31. Off
	14. Zoom in	Selection	32. Select single
	15. Zoom out		33. Select group
	16. Scroll		34. Find
	17. Next		
	18. Previous		

Figure 17: The list of 34 referents used in the study.

3.1.4 Procedure

At the start of each session, participants were asked to fill out a short survey regarding prior experience with related devices. Participants were then informed of the purpose of the study, before being *primed* [31] with a short introduction to SHMGs. This included defining SHMGs as gestures performed on the surface of the hand, from the wrist to the fingertips, using only the fingers of the same hand, without interaction with other objects or devices. Participants were allowed to use either hand. Several types of actions, *e.g.,* tap or swipe, were shown to the participant, with variations of each explained to give them a better understanding of what gestures might be considered unique. Examples of such variations are shown in Figure 19.

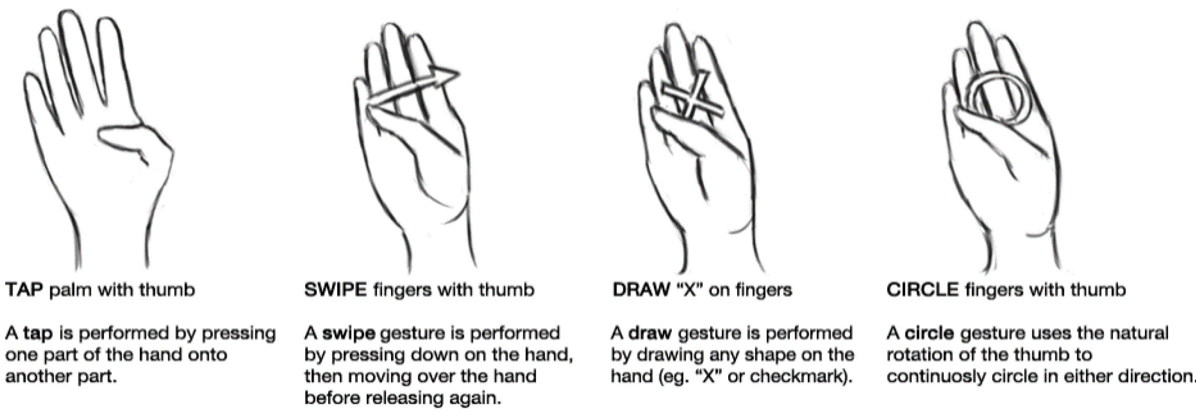

Figure 18: The four types of actions in SHMGs are illustrated and defined.

Participants were encouraged to design gestures based on preference without concern for implementation feasibility. However, participants were undoubtedly affected by previous associations with other input methods or implementations; this is further discussed in the results and analysis sections. We also specified gestures could be reused for different tasks if it made sense to participants.

Participants were presented with a list of the 34 referents and asked to design three gestures for each task, before identifying their preferred gesture for the task. Referents were always presented in the same order to participants. Based on other gesture elicitations, three gestures was a good number to apply the *production* technique to reduce legacy bias [31]. Asking for too many gestures would result in the study becoming much longer; the participants may begin experiencing either mental or physical fatigue. Each referent was demonstrated on the laptop computer, and participants were allowed to ask for clarification if required (for example, about the difference between "Move" and "Pan"). Sometimes, participants were asked to explain their choices for greater understanding of their thought process. After they completed each of the referents, we performed a semi-structured interview to elicit feedback about their experience, including potential use cases and difficulties encountered. Participants were generally enthusiastic to provide their opinion, which was encouraging for both the study and use of SHMGs.

3.2 Data and Analysis

From our 16 participants, we collected a total of 1,632 gestures (16 participants x 34 referents x 3 gestures). Appendix B contains the full data set of elicited gestures. These gestures were classified with the aforementioned methodology; this process is explained in detail in the following section. From the resulting set of gestures, we calculated agreement rates between participants and interpret them. A consensus set, as defined by Wobbrock *et al.* [59], is presented for SHMGs.

3.2.1 Classification of Gestures

In elicitation studies such as the study by Wobbrock *et al*. [59], the authors first record and transcribe each gesture, then classify them based on a set of rules. Quite often, Nielsen's *descriptive labeling* [35] and Buxton's *chunking* and *phrasing* techniques [2] are incorporated into these rules. Gestures are recorded based on their actions, rather than their semantic meanings. A gesture is then made up of one or more distinct actions. The results can then be compared and analyzed, resulting in an agreement rate, a consensus gesture set, and other general observations. When the gestures have been classified, identical gestures are grouped together. However, our approach resembles that of Piumsomboon *et al.* [39], since we group gestures that are similar rather than identical (see Appendix B for more details on classification methodology).

The primary distinction of gestures used is the type of action performed in the gesture. In hand gestures, these actions are performed by one or more fingers. We were able to categorize all the elicited gestures into four actions: *Tap, Swipe, Circle,* and *Draw.* Definitions and examples of each action were illustrated in Figure 20.

During the study, participants were asked to pick their preferred gesture after coming up with three unique gestures. In many cases, the participant would remember the action they originally performed, but mix up the exact finger(s) used. This observation is consistent with another study, in which users expressed little concern about exactly how many fingers were used in a gesture [60]. The confusion was also seen when comparing hand poses, where the fingers not used in the gesture would be bent in one variation but not the other. When reviewing the recordings, we were surprised by how often this happened. To account for this confusion in recalling gestures, we separated gestures that used two or less fingers from those with three or more fingers. This is less restrictive than matching the exact finger(s) used in each gesture, and better represents the reasoning behind gestures. Some participants commented on using one or two fingers for more precise actions (such as “Select Single”) while using three or more figures for tasks that seemed to need more space (such as “Select Multiple” or “Move”).

From the original 1,632 gestures, we isolated the 544 preferred gestures (140 unique gestures). By following the grouping approach defined above, we were able to reduce the set to 47 unique gestures. By taking the maximum consensus gesture for each referent in this set, we were left with 8 unique gestures, which represented 220/544 gestures or 40.4% of the entire set (see Appendix B for more details on the classification methodology).

3.2.2 Agreement Rate

The original agreement rate formula was developed by Wobbrock *et al*. and adopted by many gesture elicitations. Shortly prior to our research, an improved agreement rate formula was proposed by Vatavu and Wobbrock [52], which addresses some inaccuracies in the old formula. We measured agreement between participants using this new formula and the accompanying AGATe (AGreement Analysis Toolkit) software (Figure 21). The revised agreement formula is defined as

$$AR(r) = \frac{\sum_{P_i \subseteq P} \frac{1}{2}|P_i|(|P_i|-1)}{\frac{1}{2}|P|(|P|-1)}$$

where "P is the set of all proposals for referent r, $|P|$ the size of the set, and P_i subsets of identical proposals from P" [52]. The agreement rate ranges from 0 to 1, and their associated interpretations are in Figure 20.

$\mathcal{AR}(r)$ INTERVAL	PROBABILITY †	INTERPRETATION
$\leq .100$	22.9%	*low* agreement
$.100 - .300$	59.1%	*medium* agreement
$.300 - .500$	14.1%	*high* agreement
$> .500$	3.9%	*very high* agreement

† According to the probability distribution functions shown in Figure 2 and $|P| = 20$ participants.

Figure 19: Qualitative interpretations of agreement rates were proposed by Vatavu and Wobbrock, based on results from previous elicitation studies, and Cohen's guidelines for effective sizes [5].

Agreement rates ranged from 0.042 (low agreement) to 0.650 (very high agreement). The mean *AR* was 0.191 (medium agreement). The agreement rates of all referents are shown in Figure 22. Since the new formula calculated *AR* less optimistically, Vatavu *et al.* recalculated the *AR* of 18 previous studies [52]. In these studies, the average sample size was 19, while mean *AR* was 0.221.

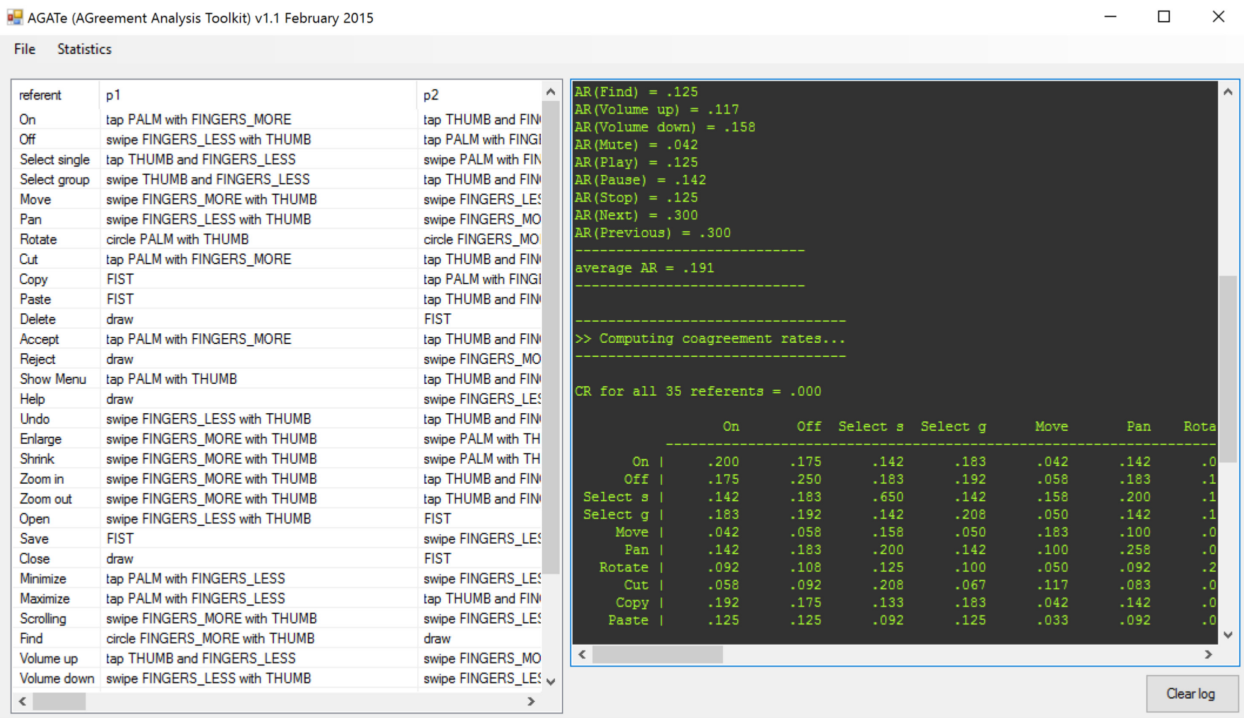

Figure 20: Vatavu and Wobbrock's AGATe Toolkit simplifies the calculation of agreement rates (AR), co-agreement rates (CR), and disagreement rates (DR).

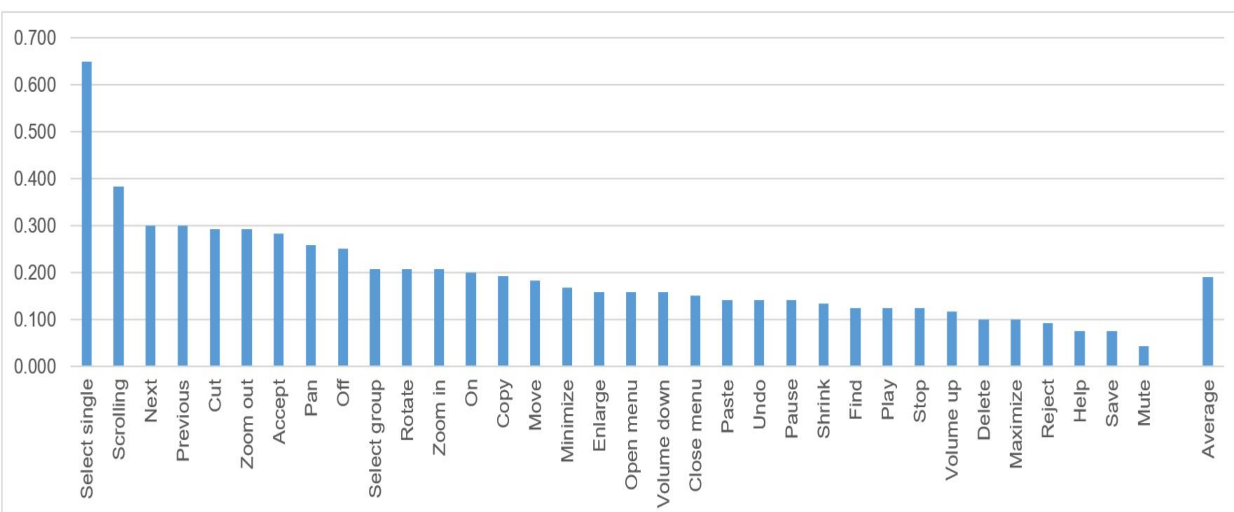

Figure 21: Agreement rates for the 34 referents.

Along with a new formula for agreement rate, Wobbrock *et al.* also introduced the *coagreement rate*, which looks at "how much agreement is shared between two referents r_1 and r_2." This is interesting because we can observe patterns previously left unnoticed. Existing work had already shown a significant relationship between gestures of dichotomous pairs [39,46,59] such as "Next"/"Previous" or "Zoom In"/"Zoom Out". Most of the pairs described were directional, where consensus gestures opposed each other directionally *(e.g.*, swipe left/right). Less focus has been placed on toggles such as "On"/"Off" or "Play"/"Pause". With the new Coagreement Rate, we not only found that "On"/"Off" (as well as "Play"/"Pause") have the same consensus gesture, but that participants who picked one gesture in r_1 often picked the same gesture in r_2. We know this since the AR for r_1 and r_2 are close to $CR(r_1,r_2)$. For example, $AR(\text{On}) = 0.200$, $AR(\text{Off}) = 0.250$, and $CR(\text{On,Off}) = 0.197$. This is different from only knowing that the same number of participants picked the consensus gesture in both referents, and suggests referents of this type should use the same gesture (see Appendix C for dichotomous pairs).

3.2.3 Consensus Gesture Set

As mentioned in Section 3.2.1, the original gesture set was reduced to 8 unique gestures. This set is rather small and even within each of the six categories of referents, there were conflicts where one gesture was preferred for several referents. This was an expected outcome, since we classified five fingers with only two categories: two fingers or less, and three fingers or more. To resolve the conflicts, we looked at each instance of the consensus gesture for each referent and identified which fingers were used most. The idea behind this resolution comes from observing the participants. While participants often mixed up the exact finger they suggested for a gesture, there was a recurring theme of choosing similar gestures for seemingly related tasks. Several participants exhibited this pattern when choosing gestures for "Cut", "Copy", and "Paste", as well as "Accept" and “Reject”. We observed a strong preference for keeping these gestures “close to each other” or “next to each other”.

Sometimes participants arbitrarily chose different fingers for a similar gesture (such as tapping any finger and the thumb together), when they had difficulty coming up with three meaningful gestures. We tried to reduce this source of randomness by taking the most used finger(s) for each consensus gesture. Very interestingly, assigning fingers with this procedure resolved all but one conflict in the consensus gesture set.

The only remaining conflict was between “Stop”, “On”, and “Off”. Since the top two preferred gestures for each of these referents were the same (make a fist, or tap the index/middle/ring ringers on the palm), we included both gestures for all three referents. We suggest using the same gesture for both “On” and “Off”, as we previously mentioned a significant coagreement rate between the two. The resulting consensus set of 16 gestures representing 34 referents in 6 categories is shown in Figure 23.

Figure 22: Consensus Gesture Set. After classifying the elicited gestures, there were 8 gestures remaining. One example is tapping the thumb with *less than two fingers.* These gestures did not specify which fingers should be used specifically. Combining these 8 gestures with finger preferences for each referent, we obtained a set of 16 consensus gestures.

3.2.4 Actions

To better understand the distribution and makeup of the gestures elicited, recall our classification method which separates gestures by actions, based on Bill Buxton's work on *Chunking* and

Phrasing [2]. When we examined the actions chosen for consensus gestures, we discovered several motifs.

Of the four action types, *Taps* were the most common (19 of 34 referents). During the think-aloud sessions, users offered some potential reasons for picking *Taps*. *Taps* were popular amongst users because of their ease with which they can be performed and their conceptual simplicity, making them easy to reproduce. Many *Tap* gestures were also preferred due to their resemblance to interaction with other devices, such as mice, trackpads, gaming devices, or remote controllers. This is apparent in the *Selection* category, where all three consensus gestures used *Taps*. A *Tap* gesture provided the precision desired when selecting a specific set of objects.

Swipes (14 of 34 referents) were frequently used when the task involved picking a value inside a continuous range, such as turning the volume up or down. In many cases they reminded users of the fluid action of sliders or radial dials. *Swipes* were also often used for tasks that were directional, such as moving something or scrolling in any direction. Of the six referents in the *Transforms* category, five made use of *Swipes.* The "Rotate" task used a *Circle* action, which was likely chosen due to the circular motion associated with rotation.

The *Draw* action appeared with six of the participants, but did not make it into the consensus set. Although drawing a question mark for "Help" or drawing an 'X' for "Close" seemed more intuitive and easier to recall, participants only resorted to the *Draw* action when experiencing difficulty devising three gestures.

Compounds gestures made up 12% of all gestures elicited, and were preferred for approximately 10% of tasks. These were mostly used for tasks that users split into smaller modules. For example, when asked to select a group of items, a participant said, "I swipe across my fingers like I am choosing the items, then I tap on my fingers to select them." In another example where a participant was asked to perform the *Save* task, the participant responded, "I have something here, then I want to make a copy here to save it."

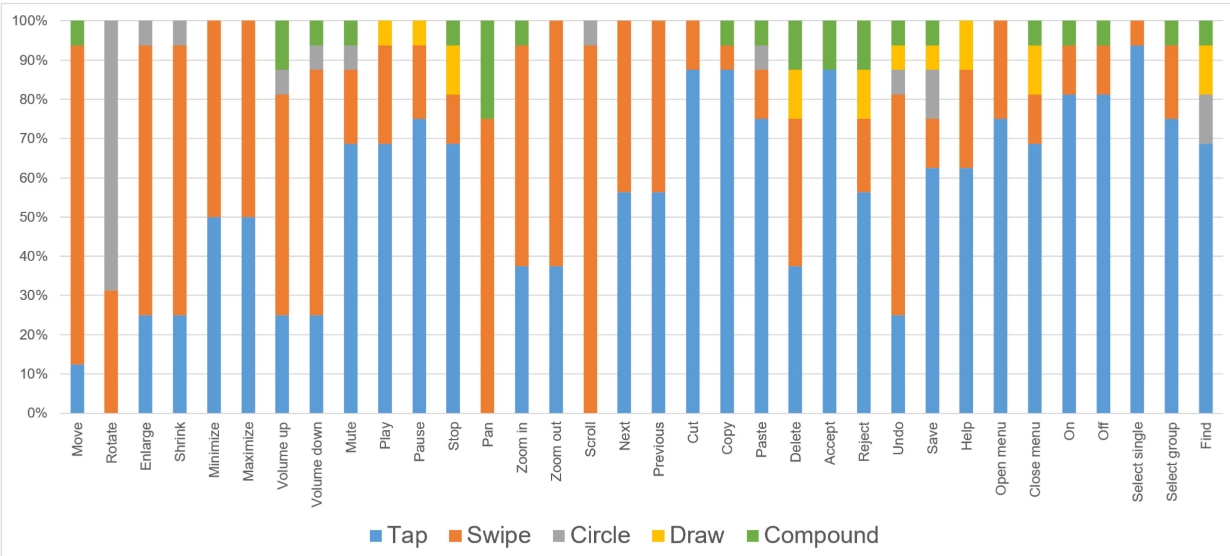

Figure 23: Distribution of action types in the preferred gesture set.

In Figure 24, we present the distribution of action types in the preferred gestures set. The preferred gestures set represents the preferred gestures of every participant, rather than just the gestures in the consensus set. By examining the graph, we can observe which actions were preferred for specific gestures. For example, we can tell that *Swipes* were preferred for dichotomous pairs (eg. "volume up" and "volume down", "enlarge" and "shrink"), which are discussed in more detail later.

3.2.5 Actors

Given the physical constraints of SHMGs where gestures are performed using only a single hand, it made sense that all gestures were performed using one or more finger(s). As motivated by Nielsen, part of designing a good gesture is to ensure the gesture is ergonomic. Knowing which fingers were most common in our data helps us to quantitatively assert which actors are most suitable for SHMGs. We can then combine qualitative observations from the study with insights from existing work to suggest reasons for some of the actors standing out as most commonly used by participants. The frequencies of each finger appearing in the gestures elicited can be seen in Figure 25.

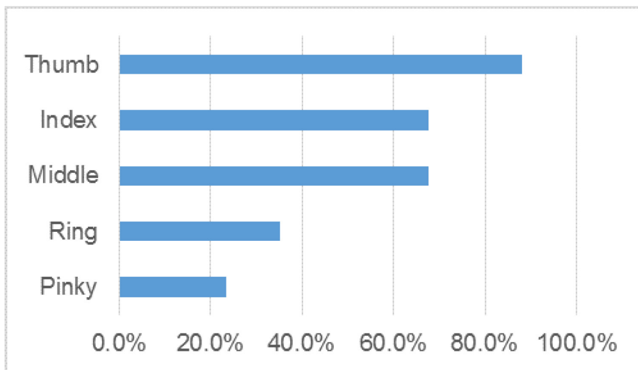

Figure 24: Frequency rates of each finger in consensus set. Sum > 100%, as multiple fingers may be used in a single gesture.

Unsurprisingly, the thumb was involved in 88% of all gestures. As explained by existing work relating hand anatomy and gestures [60], our hands are opposable through the use of our thumbs. Because of this special trait of thumbs, as well as its unique ability to rotate, the thumb can easily touch other parts of the hand, which by definition of SHMGs constitutes a gesture. Whereas other fingers have difficulty interacting with their neighbors, thumbs can touch most areas of the other fingers quite naturally. Capable of rotating, the thumb is often used for controls that involve rotation or multiple axes. For example, many gaming controls use the thumb for the D-Pad or joystick, while (two-dimensional) phone screens are often interacted with the thumb. Similarly, all the elicited *Swipe* gestures were performed with the thumb.

The preference of the index and middle fingers, when compared to the ring and pinky fingers, can also be explained by Wolf *et al.'s* summary of the anatomy of the hand [60]. Due to biomechanics and more specifically the muscles involved in moving each finger, the index finger is most suited to independent movement, followed by the middle finger. The ring finger is considered to be the least feasible, because “two muscles (*M. flexor digitorum profundus & M. flexor digitorum superficialis*) are bending synergistically the index, middle, and little finger to bring them into the palm position. In addition another muscle is responsible for stretching the ring finger (*M. extensor digitorum*), but because this muscle is also responsible for stretching the other fingers and because the ring finger has a physical connection to the middle finger (*Connexus intertendineus*), the middle finger will always move a bit in the same direction as the ring finger does.”

While the pinky finger is also able to move independently like the index finger, it was seldom used in the consensus set (2 of 34 referents). Possible explanations include the greater distance between the thumb and pinky finger, as well as the reduced strength of the pinky finger compared to the index finger. Some participants avoided using the pinky finger due to potential discomfort and fatigue.

This chapter detailed the methodology for the study, and discussed the classification of the elicited gestures. Quantitative analysis was performed on the gestures, resulting in agreement rates for each gesture as well as a consensus gesture set. Furthermore, the distribution of gestures was analyzed based on the actions performed in each gesture, and the actors (fingers) performing these gestures. In the next chapter, the quantitative analysis is combined with user feedback attained during the interviews, at the end of each elicitation. Based on both quantitative and qualitative observations, we present several themes which were prevalent during the elicitation.

Chapter Four: User Comments and Design Implications[2]

Combining the results of Chapter Three and the feedback obtained in from users at the end of each elicitation, we derived several persistent themes which were observed throughout the study. These themes are an important contribution to our third and last research goal, which is to provide guidelines for the design of SHMGs. We first present the themes, before continuing with more in-depth descriptions of each theme.

1) *Aliasing* should be used when possible, by assigning multiple gestures to the same referent.
2) *Legacy bias*, or preferences due to past experiences, should be leveraged through the use of metaphors. If no suitable metaphors exist, then abstract gestures which are physically simple to perform should be used.
3) *Meanings behind fingers and postures* can serve as metaphors as well. However, negative connotations in various contexts and cultures can lead to gestures being avoided.
4) *Dichotomous pairs*, such as “Next”/”Previous” or “Volume Up”/”Volume Down”, should be grouped and assigned opposing gestures. *State toggles*, such as “On”/”Off” or “Play”/”Pause”, should be assigned identical gestures.
5) *Details* in gestures should not be overly specific, to avoid becoming difficult to remember or recall. Detailed gestures will also be difficult perform with precision, if the interaction areas are too small (eg. tap the middle joint of a finger).
6) *Comfortable* gestures are preferred, especially for highly repetitive gestures.

---

[2] The contents of this chapter are based upon [4].

4.1 Aliasing

The average agreement rate amongst users was 0.191, which represented medium agreement between users when eliciting SHMGs. While the consensus gesture set should not be disregarded, it also cannot be proposed in whole as an optimal set of SHMGs. The agreement rates for individual gestures ranged from low agreement (0.042) to very high agreement (0.650), with 6 of 34 gestures showing low agreement, while 24 had medium agreement and 4 having high agreement. Our classification method of grouping similar gestures helped to improve the agreement scores, but it would be difficult to further group similar gestures together. Instead, it may be beneficial to group different gestures together when implementing a set of gestures. Using the *aliasing* technique suggested in other elicitation studies [46], an implementation of SHMGs may include both the first and second most frequent choice of gestures for each referent or task. As suggested by one participant, "If you have both [gestures] doing the same thing, it would be less easy to make a mistake."

The distribution of actions performed in each gesture in Section 3.2.4 makes it clear that the majority of gestures are performed with swipes or taps. On average, the most preferred action type for each referent covers 68.9% of all proposed gestures. However, by including the second more preferred action type, 91.2% of all gestures are accounted for. By implementing a combination of gesture types for each referent, such as both a swipe gesture and a tap gesture, more users can be matched with their preferred gestures.

If *aliasing* were to be used, then there will also be more conflicts of gestures being used for multiple referents. To successfully implement *aliasing*, it is necessary to determine the usage context of each referent. If the context can be identified based on the six referent categories (Section 2.1.3), then conflicts in gesture preference only have to be resolved for the referents within each category, rather than for all 34 referents. Some participants came to the same conclusion, exemplified by the following observation:

> *"I would think [that] zooming in and out, and shrinking and enlarging would be the same thing. But you would use it in a different context. You just used different words because you would label it for different contexts. But they would be the same thing."*

4.2 Previous Experiences and Legacy Bias

While the agreement rate was comparable to existing studies, we also believe previous experience of participants strongly affected our results. The significant effect of legacy bias may be due to the sheer number of interactions we perform with our hands on a daily basis, including interactions with mundane everyday objects. This motif has been previously documented [46,53], and generally led to greater agreement amongst participants. However, we found that the previous experience of our participants both positively and negatively affected agreement rates. An example where it contributed positively to agreement is the "Cut" referent, which users easily associated the task with a common symbol for scissors (tapping the index and middle fingers together). In another example where previous experience may have negatively influenced agreement rates, the proposed gestures for "Mute" included using the sign language representation of the letter "m" and also simulating the action of reaching towards the back of a handheld gaming console to reduce volume. While there are physical representations for "Mute", such as a clenched fist in music performances (a gesture that participants had little prior experience with), users drew on a large variety of other previous experiences for such actions. Regardless of whether previous experiences affected agreement rates positively or negatively, the impact of these experiences was apparent in the behavior of participants. In particular, many participants drew inspiration from electronic devices, such as gaming controllers and devices, tablets, and smartphones.

> *"I think something that mimics what we have common sense about. When I think about next, you might think the next one is always on the right hand side, as opposed to going left. So you always go on the right hand side. And then previous is like, left. And I guess a lot of my choices are affected by the technology we have now. I mean, think about it, my iPhone, your phone. It affects me. It's just easy to remember."*

Köpsel *et al.* argued that legacy bias should be leveraged to shorten the time and effort necessary to learn new ways of interaction [19]. Diverse past experiences can negatively impact agreement between users, yet sufficient exposure to similar experiences can improve agreement rates. As documented by Nebeling *et al.*, we also noticed a trend where referents which related to physical actions (such as "Cut") resulted in greater recall and agreement when metaphors were used [34]. This observation suggests that gesture designers must consider the nature of each referent, the existing metaphors, and whether these metaphors are commonly used by the expected users of the system. For referents that do not benefit from the use of metaphors, abstract gestures are more suitable as indicated by the numerous cases when users recalled specific details incorrectly.

4.3 Fingers and Postures: Their Meanings

Another topic that surfaced in other elicitation studies is the cultural meaning of various hand postures and gestures. While symbolic hand gestures have already been discussed [46,53], *e.g.*, "Help" with a beckoning gesture or "Mute" with a clenched fist, we found that users often chose specific fingers as well for a variety of reasons. Besides using the index finger for its dexterity or convenience, users frequently referred to the index finger as the pointer finger, which evoked a feeling of confidence or direction. A particularly interesting case is "Help", where one user used the pinky finger because "pinky is the weaker one, so you need more help."

SHMGs can be discrete and subtle, but we expect these gestures to be performed in both private and public spaces. As such, certain gestures may be less suitable than others and may need to be substituted for specific user groups. One user admitted that, “depending on where it is, I would not do it, just because some gestures could be bad. So I think it depends on where [I am] in public, and the culture.” Another user, when asked whether they would perform these gestures, wanted to verify the discrete nature of SHMGs: “Since they are microgestures, I could probably just do them on the side and no one would know right? I do not have to hold my hand out do I? From these concerns, it is apparent that if the gestures available are not suitable for use in the current context, then the users would simply avoid performing them at all.

4.4 Dichotomous Pairs and State Toggles

Another reason for choosing specific fingers was the motif of dichotomous pairings, and in some cases groupings of three or more gestures. As previously mentioned, dichotomous pairs often resulted in opposing gestures, such as swiping left to symbolize previous and swiping right to symbolize next. “Pairs should always go together, it is like opposite reactions. That is what I think. I want to do one way for maximize, [and] I want to do the exact opposite for minimizing so it is easy to remember.” In Figure 26, *Swipes* were shown to be preferred for “Enlarge”/”Shrink”, “Minimize”/”Maximize”, “Volume up”/”Volume down”, and “Zoom in”/”Zoom out”. As *Swipes* were heavily preferred for dichotomous pairs in the consensus set as well, we again make the recommendation to *Swipes* for these gestures.

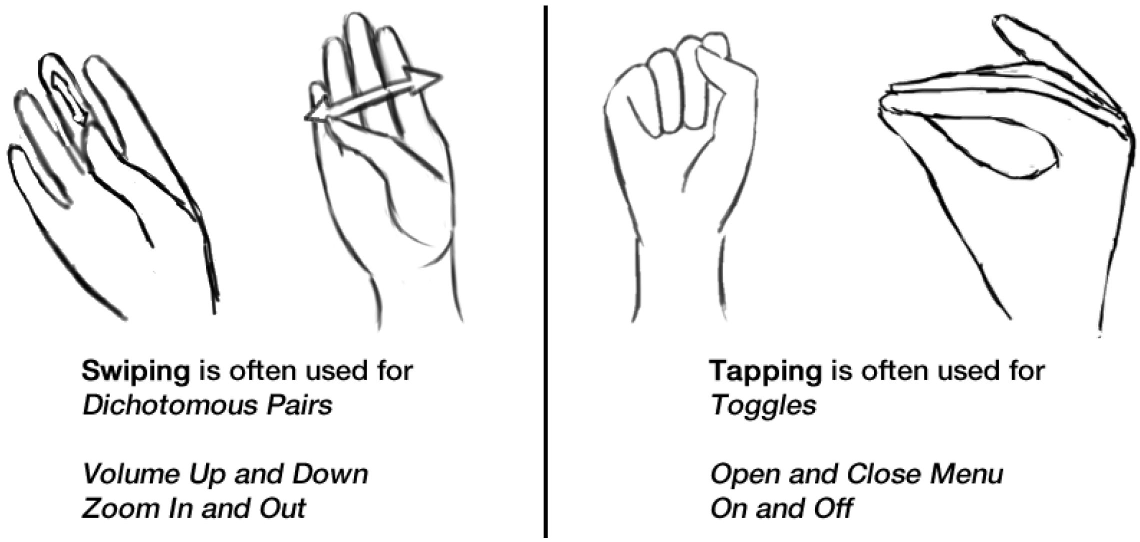

Figure 25: Examples of Swipe gestures used for dichotomous pairs (left), and Tap gestures used for toggles (right).

We also recommend using identical gestures for toggles, such as "On"/"Off". Identical gestures are more suited to toggles than opposing gestures, as we identified a unique problem with hand gestures when applying certain gestures for toggles. A good example is when some users suggested closing their fist to turn the system "On", while releasing their fist to turn the system "Off". Although the gestures are unique, the hand naturally returns to a relaxed state after tension, relating to Buxton's delimitation of atomic gestures through periods of tension and relaxation [2]. As such, performing the "On" gesture results in the "Off" gesture also being performed. This difficulty was encountered for other pairs such as "Enlarge"/"Shrink" or "Volume Up"/"Volume Down", forcing users to choose other gestures.

While all users frequently suggested some grouping of dichotomous gestures or state toggles, there was one exception which did not always follow this rule.

> *"I feel like this [accept] gesture would be fine [without an opposing gesture to reject], because it is so important; accepting a request, an email, rejecting an offer... For those things, I feel like they have to be really different, because it is not like you can fix it when*

> *you do it. You have to give a concrete answer. Whereas these ones [minimize and maximize], if I minimize it too much, I would go back and maximize it. Whereas this one it is concrete and it is done, so I would make them really different."*

4.5 Level of Detail

Given the variety of actors, actions, and interaction areas, there are technically hundreds of possible SHMGs. However, while some users went as far as using different joints to differentiate gestures, most users settled for less detail in their gestures. Many users even complained about the lack of gestures available, as one participant described: "It's very limited, (the) amount of things you can do with one hand and touch." The difficulties participants experienced in recalling gestures in detail prompted the classification method used in our study, where we grouped fingers together. When asked about their difficulties in using SHMGs, a participant suggested:

> *"Maybe trying to remember what all the gestures are…. I guess it depends on how detailed the list is, because I noticed sometimes you asked if it's the front or side [of the finger], and if it mattered. And like, if you have your finger straight or curved."*

Another participant worried "some people would be limited in the number of hand gestures they would have based on hand mobility." This was the case for one participant who could not form a clenched fist, due to having a weak grip. The dexterity of users could influence their preference of gestures.

Finally, select users were aware of variables for creating gestures but opted not to use them, as is the case when one participant used double taps instead of holds (long duration tap). The participant preferred double tapping, which felt more reassuring to them than holding a gesture for a specific duration.

4.6 Comfort

Although many different factors affected how users came up with their gestures, participants always commented on how comfortable a certain gesture was. In many cases, they even selected their preferred gestures for referents based on comfort. As discussed in Section 3.2.5, the biological anatomy of the hand made some gestures more comfortable and easily performed than others.

> *"I would probably just slide it on one finger. Usually the middle finger, because it is more comfortable. And then slide the other way if I want to rotate the other way. I pick the middle finger gesture, because it would be tiring to do the other gesture on the palm, with my thumb."*

Reducing the distance between the acting finger and the interaction area also contributed towards greater comfort. Participants often described the distance within gestures using terms such as "close", "just right there", or "easy to reach."

Although comfort was very important for selecting gestures, most participants expressed little concern for potential fatigue using these gestures. After some consideration, a participant said they "do not think there will be fatigue unless you are doing it in an exaggerated manner, or doing too much of something. If you are just tapping or swiping, I think it should be fine."

Chapter Five: Study Limitations[3]

As the very first elicitation study on SHGMs, there were a number of limitations which we discovered. The first limitation concerns the technology and definition of SHMGs. The second limitation is relevant to the elicitation methodology. The third and fourth limitations concern the study in general, discussing the demographics of the participants as well as the context of the study.

5.1 Redefining SHMGs

5.1.1 Additional Variables

Due to the perceived limitation of gesture variety, users reported two interesting variables that they could potentially control in addition to the suggestions we made. First, they suggested varying the speed at which a gesture is performed. Performing a gesture slowly was perceived to offer finer adjustment, such as when performing the "Enlarge" or "Shrink" tasks. The second variable used varying forces while performing gestures such as closing a hand harder to perform "Stop" instead of "Pause". These variables may enable a larger vocabulary of natural gestures, provided the speed and force can be detected reliably. These user suggestions were made during an interview at the end of the study, but no such gestures were chosen by participants in the elicitation part.

5.1.2 Spatial Tracking

As seen with the additional variables proposed by users, our current definition of SHMGs may not match the expectations of all users. This is, after all, the motivation behind consulting users in our elicitation study. Participants were asked to comment on the feasibility of SHMGs as well as the study itself, and all participants commented on not being able to use mid-air 3D or spatial gestures. For example, participants asked if they could perform "Move" by tapping the thumb and index

[3] The contents of this chapter are based upon [4].

fingers together, before moving the whole hand in mid-air. One participant argued that "it would be nice if you had more than just your hand [for gesturing]. I feel like there is only so much I can do, that are not all the same." Later on, they repeated the suggestion:

> *"I know they have this armband [that can do spatial tracking]. Maybe incorporate that and the microgestures together, so you can have more buttons, or more that you can do. I think it would be easier to use with [spatial tracking], because then you would have more choices."*

Although we defined SHMGs as gestures performed on the hand from the wrist to the fingertips, many users would have liked the option of using spatial tracking of the arm itself as well. While larger arm movements may not be suitable for discrete microgestures in public spaces, users frequently proposed small movements or rotations of the arm. This happened despite users being informed during *priming* that such spatial gestures did not fit our criteria, suggesting the desire and possible need for spatial recognition.

5.2 Legacy Bias in Elicitations

As documented by existing literature [46,53], legacy bias may have a significant effect on results in an elicitation study. Morris *et al.* proposed the *priming, pairing,* and *production* techniques for reducing legacy bias. While these techniques may help, they are unable to completely mitigate the effects of legacy bias.

5.2.1 Priming

Priming familiarizes the participants with the system or interface, so that they are more likely to embrace the new system, rather than fall back on older "safe" experiences when proposing gestures. This usually involves an explanation and demonstration of the system, before the elicitation begins. *Priming* certainly benefited the users, as they often thought back to the demonstrations to remember what gestures can be performed. However, many users asked questions during the elicitation to clarify any newly discovered uncertainties about the system. As proposed by a participant:

> *"I think you should draw pictures of hands, and of things you can do, as opposed to showing them. It is just nice to have a reference, as opposed to you telling them [during the study], because that is going to affect what they think as you tell them."*

Indeed, *priming* may be extended to become a reference for participants both *before* and *during* the study. Having these neutral resources available to the participants can give them more confidence to try different gestures, without introducing experimenter expectancy effects.

5.2.2 Pairing

Pairing participants together is another way to reduce legacy bias [31]. With previous user experience and legacy bias having both positive and negative effects on agreement rates, *pairing* may be useful as a means to generate more optimal gestures. In the situation where a single user might run out of ideas and therefore offer arbitrary gestures as their second or third choice, having a partner may help foster additional ideas. When users pick gestures based on personal and unique experiences, a partner would be able to question the generalizability of such a gesture in a consensus set. While pairing could help to improve the results, a dominant participant could also bias the collaborative efforts. Additionally, the significantly increased time requirements, as a result of having more participants, often makes *pairing* impractical for elicitation studies. For this first elicitation of SHMGs, we opted for not using the *pairing* technique, since we would not have a reference set of data (without *pairing*) to compare our results with.

5.2.3 Production

Production can offset legacy bias by requiring participants to propose multiple gestures for each referent. Even should they propose a gesture influenced by legacy bias, they might still propose subsequent gestures that are less affected. However, we encountered the same problems mentioned by Morris *et al.* when applying *production* [31]. That is, there is no way to determine the optimal amount of gestures each user should propose for each referent. In 55% of all cases, users did indeed choose their second or third gestures as their preferred gesture. When later asked why they did not propose gestures that seemed obvious to the researcher, users often replied, "I didn't even think of that!" However, other users benefited less from *production*: "I already have a gesture in mind, so thinking of three different ones makes me start grasping for straws, because I already have a solid idea of what I would do."

5.3 Study Participants

For the study in Chapter Three, participants were recruited through email lists and word of mouth. Participants were not screened based on their demographics or experiences with gesture input technology. As a result, the participants were relatively young, with 16 participants averaging 22 years old. These demographics are quite common for elicitation studies, when compared to the 18 elicitations discussed by Vatavu *et al.* [52]. In 17 of the 18 studies, the number of participants ranged from 12-22, with 35 participants in the remaining study. Of the 12 studies which reported average ages, 11 studies had an average age between 20 and 30 years old.

Although our demographics are consistent with existing elicitation studies, these demographics may still pose a significantly impact on the results. As such, it would be difficult to generalize these findings for other age groups, without further studies. It would be informative to conduct further studies, either to compare with different age groups, or as an attempt to repeat the results of this study.

5.4 Specific Domains

It would also by worthy to investigate user preferences in more specific domains suited to SHMGs, such as while in public transit or while performing a primary task in parallel. While the inherent nature of SHMGs makes them less susceptible to factors which create social awkwardness [42], developing generic principles that apply universally to all contexts remains a challenge [35]. Further context-specific studies may reveal subtle factors specific to SHMGs that affect the gestures preferred by users.

Chapter Six: Future Work and Conclusion

We recognized the potential of single-hand microgestures (SHMGs) in ubiquitous computing, amidst current technological developments. Chapter One began with an introduction to some important concepts discussed in this thesis, and stated the research goals and questions. In adherence to the research goals, Chapter Two first presented an overview of the state of SHMG technology and research. It then thoroughly described the elicitation methodology, which is suitable for eliciting SHMGs. Chapter Three detailed the elicitation study with end users, where we recorded a set of 1,680 gestures. We presented our data with quantitative findings, including agreement rate calculations and frequency statistics. A consensus set of SHMGs was also derived from the elicited gestures. Chapter Four then contained a number of themes which form the guidelines for designing SHMGs. The proposed guidelines address the last research question, in providing future designers and implementers of SHMGs with a solid foundation.

6.1 Conclusion

Gestural input is becoming increasingly common, with devices such as the Microsoft Kinect [27] or Myo [51] armband available to the average consumer. As devices continue to shrink in size, the interaction space on the device is also reduced. With less space on devices for interaction areas or controls, gestural input has become increasingly appealing as an alternative. Recent advances such as CyclopsRing [3] or BackHand [25] have made it possible to recognize discrete and localized gestures, leading to greater interest in hand gestures such as SHMGs. Despite the potential for SHMGs to be used in ubiquitous and portable computing, there is no prior user elicitation of SHMGs. To maximize the potential of SHMGs on user devices, it is important to understand user preferences for designing SHMGs, and to design gestures that conform to the user rather than to technological limitations.

This thesis highlighted the lack of user involvement in the design of SHMG, and presented an exploration into user preferences through a gesture elicitation study. A quantitative analysis measured user preferences for the actions performed in each gesture, the fingers performing each action, and the part of the hand acted upon. An agreement rate was calculated for each referent, to determine user consensus in gesture choices. Finally, the quantitative results were further interpreted along with user comments, to derive a set of guidelines for designing optimal SHMGs.

This thesis summarizes previous work related to SHMGs, so that future designers and implementers of SHMGs can quickly grasp the current state of SHMG technology and research. Following the guidelines presented in this thesis, these designers and implementers can then produce SHMGs that are truly suited to end-users.

6.2 Future Work

The limitations in Chapter Five raise new questions for the elicitation of SHMGs, which would be interesting to pursue in future work. Studies which include additional variables or capabilities, such as spatial tracking of the arm, can determine how important such features are to users. While we do not expect significantly higher agreement rates when introducing greater variation, the newly elicited gestures may be more natural for users. If the new gestures are indeed more preferable to users, then they may also be easier to recall for users [35,48,19]. It is possible to do such a subsequent study in the near future, since very little change has to be made to the study from Chapter Three.

Another short-term variation of this study could address the context-dependency of SHMGs. Using the results from this study as a baseline, further studies in specific contexts can determine whether SHMGs are more suitable in specific contexts, which may result in an easier adoption of SHMGs.

In the long-term, the elicitation methodology itself can be adapted and modified in an attempt to improve agreement rates. For example, based on user feedback, the references which comprise the *priming* technique could be made available *during* the study, rather than only *before* it begins. It would also be interesting to repeat this study using the *pairing* technique. Given the significant effect of legacy bias observed in the study, pairing could help to develop a more universal set of SHMGs. Participants in the subsequent study could also be provided with the consensus gestures of this study, after they have proposed their own gestures. This would make the *production* technique more effective, as the participants are less likely to select a poor gesture, due to having overlooked other potentially better gestures. It would also help to verify if the results of this study are repeatable to a reasonable degree.

References

1 Angelini, Leonardo, Carrino, Francesco, Carrino, Stefano et al. Gesturing on the Steering Wheel; a User-eliciated taxonomy. In *6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '14)* (New York, NY, USA 2014), Article 31, 8 pages.

2 Buxton, William A. S. Chunking and phrasing and the design of human-computer dialogues. In Baecker, Ronald M. et al., eds., *Human-Computer Interaction*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1995.

3 Chan, Liwei, Chen, Yi-Ling, Hsieh, Chi-Hao, Liang, Rong-Hao, and Chen, Bing-Yu. CyclopsRing: Enabling Whole-Hand and Context-Aware Interactions Through a Fisheye Ring. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)* (New York, NY, USA 2015), ACM, 549-556.

4 Chan, Edwin, Seyed, Teddy, Stuerzlinger, Wolfgang, Yang, Xing-Dong, and Maurer, Frank. User Elicitation on Single-hand Microgestures. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)* (New York, NY, USA 2016), ACM, 3403-3414.

5 Cohen, Jacob. A power primer. *Psychological Bulletin*, 112, 1 (July 1992), 155-159.

6 Creative Labs. *Creative Senz3D*. www.creative.com/senz3d, March, 2017.

7 Dahlbäck, Nils, Jönsson, Arne, and Ahrenberg, Lars. Wizard of Oz studies: why and how. In *Proceedings of the 1st international conference on Intelligent user interfaces (IUI '93)* (New York, NY, USA 1993), ACM, 193-200.

8 Defanti, Thomas and Sandin, Daniel. *Final Report to the National Endowment of the Arts*. University of Illinois, Chicago, Illinois, 1977.

9 Dipietro, Laura, Sabatini, Angelo M., and Dario, Paolo. A Survey of Glove-Based Systems and Their Applications. *Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38, 4 (2008), 461-482.

10 Dong, Haiwei, Danesh, Ali, Figueroa, Nadia, and Saddik, Abdulmotaleb El. An Elicitation Study on Gesture Preferences and Memorability Toward a Practical Hand-Gesture Vocabulary for Smart Televisions. *IEEE Access*, 3 (May 13, 2015), 543-555.

11 Furnas, George W., Landauer, Thomas K., Gomez, Louis M., and Dumais, Susan T. The vocabulary problem in human-system communication. *Communications of the ACM*, 30, 11 (November 1987), 964-971.

12 Goldin-Meadow, Susan. Widening the lens: what the manual modality reveals about language, learning and cognition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369, 1651 (September 2014).

13 Grimes, Gary. *Digital data entry glove interface device*. Bell Telephone Laboratories, Murray Hill, New Jersey, US, 1983.

14 Groenewald, Celeste, Anslow, Craig, Islam, Junayed, Rooney, Chris, and Wong., and William. Understanding Mid-Air Hand Gestures for Interactive Surfaces and Displays: A Systematic Literature Review. In *Proceedings of the British Human Computer Interaction Conference (BritishHCI)* (New York, NY, USA 2016), ACM.

15 Harrison, Chris, Benko, Hrvoje, and Wilson, Andrew D. OmniTouch: wearable multitouch interaction everywhere. In *Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11)* (New York, NY, USA 2011), ACM, 441-450.

16 Harrison, Chris, Tan, Desney, and Morris, Dan. Skinput: appropriating the body as an input surface. In *SIGCHI Conference on Human* (New York, NY, USA 2010), 453-462.

17 Hutchins, Edwin L., Hollan, James D., and Norman, Donald A. Direct manipulation interfaces. *Human-Computer Interaction*, 1, 4 (December 1985), 311-338.

18 Kim, David, Hilliges, Otmar, Izadi, Shahram, Butler, Alex D., Chen, Jiawen, Oikonomidis, Iason, and Olivier, Patrick. Digits: freehand 3D interactions anywhere using a wrist-worn gloveless sensor. In *25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)* (New York, NY, USA 2012), ACM, 167-176.

19 Köpsel, Anne and Bubalo, Nikola. Benefiting from legacy bias. *interactions*, 22, 5 (August 2015), 44-47.

20 Krauss, Robert M., Dushay, Robert A., Chen, Yihsiu, and Rauscher, Frances. The Communicative Value of Conversational Hand Gesture. *Journal of Experimental Social Psychology*, 31, 6 (1995), 533-552.

21 Kurtenbach, Gord and Hulteen, Eric. Gestures in Human-Computer Communication. In *The Art and Science of Inerface Design*. Addison-Wesley Publishing Company, Reading, 1990.

22 LaViola Jr., Joseph J. *A Survey of Hand Posture and Gesture Recognition Techniques and Technology*. Brown University, Providence, RI, 1999.

23 Leap Motion. *Leap Motion*. https://www.leapmotion.com/, Online, 2017, March.

24 Liang, Rong-Hao, Lin, Shu-Yang, Su, Chao-Huai, and KaiYin. SonarWatch: appropriating the forearm as a slider bar. In *SIGGRAPH Asia 2011 Emerging Technologies (SA '11)* (New York, NY, USA 2011), ACM, Article 5.

25 Lin, Jhe-Wei, Wang, Chiuan, Huang, Yi Yao, Chou, Kuan-Ting, Chen, Hsuan-Yu, Tseng, Wei-Luan, and Chen, Mike Y. BackHand: Sensing Hand Gestures via Back of the Hand. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)* (New York, NY, USA 2015), ACM, 557-564.

26 Loclair, Christian, Gustafson, Sean, and Baudisch, Patrick. PinchWatch: a wearable device for one-handed microinteractions. *MobileChi*, 10 (2010).

27 Microsoft. *Microsoft Kinect*. http://www.microsoft.com/enus/kinectforwindows/, Online, 2017, March.

28 Microsoft. *Xbox*. http://www.xbox.com/en-ca/, Online, 2017, March.

29 Mistry, Pranav and Maes, Pattie. SixthSense: a wearable gestural interface. *ACM SIGGRAPH ASIA 2009 Sketches (SIGGRAPH ASIA '09)* (2009), Article 11.

30 Morris, Meredith Ringel. Web on the wall: insights from a multimodal interaction elicitation study. In *2012 ACm International Conferenceo n Interactive Tabletops and Surfaces (ITS '12)* (New York, NY, USA 2012), ACM.

31 Morris, Meredith Ringel, Danielescu, Andreea, Drucker, Steven, Fisher, Danyel, Lee, Bongshin, schraefel, m. c., and Wobbrock, Jacob O. Reducing legacy bias in gesture elicitation studies. *interactions*, 21, 3 (2014), 40-45.

32 Morris, Meredith Ringel, Wobbrock, Jacob O., and Wilson, Andrew D. Understanding users' prefernces for surface gestures. In *Graphics Interface 2010 (GI '10)* (Toronto, ON, Canada 2010), Canadian Information Processing Society, 261-268.

33 Nacenta, Miguel A., Kamber, Yemliha, Qiang, Yizhou, and Kristensson, Per Ola. Memorability of pre-designed and user-defined gesture sets. In *SIGCHI Conference on Human Factors in Computing Systems (CHI '13)* (New York, NY, USA 2013), ACM, 1099-1108.

34 Nebeling, Michael, Huber, Alexander, Ott, David, and Norrie, Moira C. Web on the Wall Reloaded: Implementation, Replication and Refinement of User-Defined Interaction Sets. In *Ninth ACM International Conference on Interactive Tabletops and Surfaces (ITS '14)* (New York, NY, USA 2014), ACM, 15-24.

35 Nielsen, Michael, Störring, Moritz, Moeslund, Thomas B., and Granum, Erik. A Procedure for Devloping Intuitive and Ergonomic Gesture Interfaces for HCI. In Camurri, Antonio and Volpe, Gualtiero, eds., *Gesture-Based Communication in Human-Computer Interaction*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

36 Ogata, Masa, Sugiura, Yuta, Makino, Yasutoshi, Inami, Masahiko, and Imai, Michita. SenSkin: adapting skin as a soft interface. (New York, NY, USA 2013), ACM, 539-544.

37 Olberding, Simon, Yeo, Kian Peen, Nanayakkara, Suranga, and Steimle, Jurgen. AugmentedForearm: exploring the design space of a display-enhanced forearm. In *Proceedings of the 4th Augmented Human International Conference (AH '13)* (New York, NY, USA 2013), ACM, 9-12.

38 Ono, Ryosuke, Yoshimoto, Shunsuke, and Sato, Kosuke. Palm+Act: operation by visually captured 3D force on palm. In *SIGGRAPH Asia 2013 Emerging Technologies (SA '13)* (New York, NY, USA 2013), ACM, Article 14.

39 Piumsomboon, Thammathip, Clark, Adrian, Billinghurts, Mark, and Cockburn, Andy. User-defined gestures for augmented reality. In *CHI '13 Extended Abstracts on Human Factor in Computing Systems (CHI EA'13)* (New York, NY, USA 2013), ACM, 955-960.

40 Rekimoto, Jun. GestureWrist and GesturePad: Unobtrusive Wearable Interaction Devices. In *5th IEEE International Symposium on Wearable Cmputers (ISWC '01)* (Washington, DC, USA 2001), 21-.

41 Rico, Julie and Brewster, Stephen. Usable gestures for mobile interfaces: evaluating social acceptability. In *SIGCHI Conference on Human Factors in Computing Systems (CHI '10)* (New York, NY, USA 2010), ACM.

42 Rico, Julie and Brewster, Stephen. Usable gestures for mobile interfaces: evaluating social acceptability. In *SIGCHI Conference on Human Factors in Computing Systems (CHI '10)* (New York, NY, USA 2010), ACM, 887-896.

43 Ruiz, Gustavo, Rovelo, Alberto, Vanacken, Davy, Lutten, Kris, Abad, Francisco, and Camahort, Emilio. Multi-viewer gesture-based interaction for omni-directional video. In *SIGCHI Conference on Human Factors in Computing Systems (CHI '14)* (New York, NY, USA 2014), ACM.

44 Saponas, T. Scott, Tan, Desney S., Morris, Dan, Balakrishnan, Ravin, Turner, Jim, and Landay, James A. Enabling always-available input with muscle-computer interfaces. In *22nd*

Annual ACM Symposium on User Interface Software and Technology (UIST '09) (New York, NY, USA 2009), ACM, 167-176.

45 Seyed, Teddy, Azazi, Alaa, Chan, Edwin, Wang, Yuxi, and Maurer, Frank. SoD-Toolkit: A Toolkit for Interactively Prototyping and Developing Multi-Sensor, Multi-Device Environments. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces (ITS '15)* (New York, NY, USA 2015), ACM, 171-180.

46 Seyed, Teddy, Burns, Chris, Sousa, Mario Costa, Maurer, Frank, and Tang, Anthony. Eliciting usable gestures for multi-display environments. In *2012 ACM International Conference on Interactive Tabletops and Surfaces (ITS '12)* (New York, NY, USA 2012), ACM, 41-50.

47 Starner, Thad, Weaver, James, and Pentland, Alex. A wearable computer based American sign language recognizer. In *Digest of Papers. First International Symposium on Wearable Computers* (Cambridge, MA, USA 1997), IEEE, 130-137.

48 Stevanoni, Elizabeth and Salmon, Karen. Giving Memory a Hand: Instructing Children to Gesture Enhances their Event Recall. *Journal of Nonverbal Behavior*, 29, 4 (December 2005), 217-233.

49 Sturman, David J. and Zeltzer, David. A Survey of Glove-based Input. *IEEE Computer Graphics and Applications*, 14, 1 (1994), 30-39.

50 Takemura, Kentaro, Ito, Akihiro, Takamatsu, Jun, and Ogasawara, Tsukasa. Active bone-conducted sound sensing for wearable interfaces. In *Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology (UIST '11 Adjunct)* (New York, NY, USA 2011), ACM, 53-54.

51 Thalmic Labs. *Myo Gesture Control Armband*. https://www.myo.com/, Online, 2017, March.

52 Vatavu, Radu-Daniel and Wobbrock, Jacob O. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *33rd Annual ACm Conference on HUman Factors in Computing Systems (CHI '15)* (New York, NY, USA 2015), ACM, 1325-1334.

53 Vatavu, Radu-Daniel and Zaiti, Ionut-Alexandru. Leap gestures for TV: insights from an elicitation study. In *2014 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '14)* (New York, NY, USA 2014), ACM, 131-138.

54 Vicon. *Vicon Camera Systems*. https://www.vicon.com/products/camera-systems, Online, 2017, March.

55 Vogel, Daniel and Baudisch, Patrick. Shift: a technique for operating pen-based interfaces using touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)* (New York, NY, USA 2007), ACM.

56 Weigel, Martin, Mehta, Vikram, and Steimle, Jürgen. More than touch: understanding how people use skin as an input surface for mobile computing. In *SIGCHI Conference on Human Factors in Computing Systems (CHI '14)* (New York, NY, USA 2014), ACM, 179-188.

57 Wigdor, Daniel, Forlines, Clifton, Baudisch, Patrick, Barnwell, John, and Shen, Chia. Lucid touch: a see-through mobile device. In *Proceedings of the 20th annual ACM symposium on User interface software and technology (UIST '07)* (New York, NY, USA 2007), ACM.

58 Wobbrock, Jacob O., Aung, Htet Htet, Rothrock, Brandon, and Myers, Brad A. Maximizing the guessability of symbolic input. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)* (New York, NY, USA 2005), ACM, 1869-1872.

59 Wobbrock, Jacob O., Morris, Meredith Ringel, and Wilson, Andrew D. User-defined gestures for surface computing. In *SIGCHI Conference on HUman Factors in Computing Systems (CHI '09)* (New York, NY, USA 2009), ACM, 1083-1092.

60 Wolf, Katrin, Naumann, Anja, Rohs, Michael, and Müller, Jörg. Taxonomy of microinteractions: defining microgestures based on ergonomic and scenario-dependent requirements. In Campos, Pedro et al., eds., *Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction - Volume Part 1 (INTERACT '11)*. Springer-Verlag, Berlin, Heidelberg, 2011.

61 Wolf, Katrin, Schleicher, Robert, Kratz, Sven, and Rohs, Michael. Tickle: a surface-independent interaction technique for grasp interfaces. In *7th International Conference on Tangible, Embedded and Embodied Interaction (TEI '13)* (New York, NY, USA 2013), ACM, 185-192.

62 Wu, Ying and Huang, Thomas S. Vision-based gesture recognition: A review. In *Gesture-based Communication in Human-Computer Interaction. Springer Berlin Heidelberg* (New York, NY, USA 1999), ACM, 103-115.

63 Wu, Huiyue and Wang, Jianmin. User-Defined Body Gestures for TV-based Applications. In *2012 Fourth International Conference on Digital Home* (Guangzhou, China 2012), IEEE, 415-420.

64 Yang, Xing-Dong, Grossman, Tovi, Wigdor, Daniel, and Fitzmaurice, George. Magic finger: always-available input through finger instrumentation. In *25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)* (New York, NY, USA 2012), ACM, 147-156.

65 Zhang, Yang, Zhou, Junhan, Laput, Gierad, and Harrison, Chris. SkinTrack: Using the Body as an Electrical Waveguide for Continuous Finger Tracking on the Skin. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)* (New York, NY, USA 2016), ACM, 1491-1503.

66 Zimmerman, Thomas G., Lanier, Jaron, Blanchard, Chuck, Bryson, Steve, and Harvill, Young. A Hand Gesture Interface Device. In *Proceedings of CHI+GI'87 Human Factors in Computing Systems and Graphics Interface* (New York, NY, USA 1987), ACM, 189-192.

Appendix A: Pre Study Questionnaire Results– Elicitation Study

	Total	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16
Experience with gesture devices (eg. Kinect, Leap, Myo)	*25%*	N	N	N	Y	N	N	N	N	Y	N	Y	N	Y	N	N	N
Experience with touch gestures	*100%*	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Familiarity with computers	*4.94*	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Familiarity with smartphones	*4.75*	5	5	5	5	5	5	5	5	5	5	5	2	5	5	5	4
Familiarity with tablets	*4.06*	4	5	4	5	4	3	5	3	5	5	4	5	5	3	3	2
Familiarity with game consoles	*3.06*	3	5	5	3	3	3	2	1	4	1	1	3	5	4	3	3

Appendix B: Data: Elicited Gestures

referent	p1	p2	p3	p4
On	tap P with F_M	tap T and F_L	fist	fist + tap F_L with T
Off	swipe F_L with T	tap P with F_M	fist	fist + tap F_L with T
Select single	tap T and F_L	swipe P with F_L	tap F_M with T	tap P with T
Select group	swipe T and F_L	tap T and F_L	tap T and F_M	fist
Move	swipe F_M with T	swipe F_L with T	swipe P with T	tap P with T + swipe P with T
Pan	swipe F_L with T	swipe F_M with T	fist + swipe F_M with T	swipe P with T
Rotate	circle P with T	circle F_M with T	swipe F_L with T	circle F_L with T
Cut	tap P with F_M	tap T and F_L	tap F_L	swipe F_L with T
Copy	fist	tap P with F_M	tap T and F_L	fist + swipe F_M with T
Paste	fist	tap T and F_M	tap T and F_L	circle F_L with T
Delete	draw	fist	swipe F_L with T	tap T and F_L
Accept	tap P with F_M	tap T and F_L	fist	fist + swipe F_L with T
Reject	draw	swipe F_M with T	tap P with F_M	tap T and F_L
Help	draw	swipe F_L with T	tap P with F_L	tap T and F_M
Undo	swipe F_L with T	tap T and F_L	swipe F_M with T	circle F_M with T
Enlarge	swipe F_M with T	swipe P with T	swipe T with F_L	tap T and F_M
Shrink	swipe F_M with T	swipe P with T	tap T and F_M	swipe F_L with T
Zoom in	swipe F_M with T	tap T and F_L	swipe F_L with T	fist
Zoom out	swipe F_M with T	tap T and F_L	swipe F_L with T	tap P with T
Open menu	swipe F_L with T	fist	tap T and F_L	swipe T with F_L
Close menu	draw	fist	tap T and F_M	tap T and F_L
Save	fist	swipe F_L with T	tap T and F_L	tap T and F_M
Minimize	tap P with F_L	swipe F_L with T	fist	swipe F_M with T
Maximize	tap P with F_L	tap T and F_L	swipe F_L with T	fist
Scrolling	swipe F_M with T	swipe F_L with T	circle F_L with T	swipe T with F_L
Find	circle F_M with T	draw	tap P with F_L	tap P with F_M
Volume up	tap T and F_L	swipe F_M with T	swipe F_L with T	circle F_L with T
Volume down	swipe F_L with T	swipe F_L with T	tap T and F_L	swipe F_M with T
Mute	swipe F_M with T	tap F_M with T	fist	tap T and F_L
Play	swipe F_L with T	tap P with F_L	tap T and F_M	tap T and F_L
Pause	swipe F_L with T	tap P with F_L	tap T and F_M	tap T and F_L
Stop	draw	tap T and F_L	tap P with F_M	fist
Next	swipe F_M with T	tap T and F_L	swipe F_L with T	tap P with F_L
Previous	swipe F_M with T	tap T and F_L	swipe F_L with T	tap P with F_L

T = Thumb; F_L = one or two fingers; F_M = three or more fingers; P = Palm

referent	p5	p6	p7
On	swipe F_L with T	swipe F_L with T	tap P with T
Off	tap P with T	tap T and F_L	tap T and F_M
Select single	tap T and F_L	tap T and F_L	tap T and F_L
Select group	swipe F_M with T + tap P with T	swipe P with T	tap P with T
Move	swipe T with F_L	swipe F_L with F_L	tap T and F_L
Pan	tap + swipe F_L with T	tap F_M with T + swipe F_M with T	tap P with T + tap P with T
Rotate	swipe F_M with T	circle P with T	circle F_M with T
Cut	swipe F_M with T	tap T and F_L	tap F_L
Copy	swipe P with T	tap P with F_L	tap P with T
Paste	circle P with T + tap P with T	swipe F_L with T	swipe F_M with T
Delete	swipe F_M with T	tap F_L with F_L	tap P with T + swipe P with T
Accept	fist + tap F_L with T	tap P with T	tap P with F_M
Reject	fist + swipe T with F_M	fist + swipe T with F_L	swipe F_L with T
Help	tap T and F_L	fist	swipe F_M with T
Undo	draw	fist	fist + tap F_L with T
Enlarge	swipe F_L with T	circle F_M with T	swipe F_M with T
Shrink	circle F_M with T	swipe T with F_L	tap P with F_M
Zoom in	tap + swipe F_L with T	tap P with T	tap F_L
Zoom out	tap T and F_M	swipe F_M with T	tap T and F_L
Open menu	swipe T with F_M	tap P with F_M	tap P with T
Close menu	swipe F_L with T	swipe F_M with T	tap F_L+ tap F_M
Save	circle F_L with T	circle F_M with T	draw
Minimize	swipe P with T	swipe T with F_L	tap and swipe F_L with T
Maximize	swipe F_M with T	swipe P with T	swipe T with F_L
Scrolling	swipe F_M with T	swipe F_M with T	swipe F_M with T
Find	tap T and F_M	tap T and F_L	circle F_L with T + tap T and F_L
Volume up	fist + swipe F_M with T	swipe F_L with F_L	tap F_L with F_L + swipe F_L with T
Volume down	circle F_L with T	swipe F_L with F_L	tap F_L with F_L + swipe F_L with T
Mute	circle F_L with T	swipe F_L with T	tap P with F_M
Play	draw	swipe F_M with T	tap P with F_M
Pause	draw	swipe F_M with T	tap P with F_M
Stop	swipe F_L with T	swipe F_M with T	tap F_M with F_M + tap F_L with T
Next	tap P with T	swipe F_M with T	swipe F_M with T
Previous	tap P with T	swipe F_M with T	swipe F_M with T

T = Thumb; F_L = one or two fingers; F_M = three or more fingers; P = Palm

referent	p8	p9	p10	p11
On	tap T and F_M	tap P with F_M	tap T and F_L	fist
Off	swipe F_L with T	tap P with F_M	fist	fist
Select single	tap T and F_L	tap T and F_L	tap T and F_L	tap T and F_L
Select group	swipe T and F_L	tap T and F_L	tap T and F_L	tap T and F_M
Move	swipe F_M with T	swipe F_M with T	swipe F_M with T	swipe F_M with T
Pan	swipe F_L with T	swipe F_L with T	swipe F_M with T	swipe F_M with T
Rotate	circle F_M with T	swipe F_L with T	swipe F_L with T	swipe F_L with T
Cut	tap F_L	tap F_L	tap F_L	tap F_L
Copy	tap T and F_M	fist	tap P with F_M	tap T and F_L
Paste	tap P with T	tap F_L	fist	tap T and F_M
Delete	tap T and F_L+ tap P with F_L	draw	fist	swipe F_L with T
Accept	tap P with F_M	tap P with F_M	tap T and F_L	tap T and F_L
Reject	tap P with F_L	tap T and F_M	draw	swipe F_M with T
Help	swipe T with F_L	tap P with F_L	tap P with T	draw
Undo	swipe P with T	tap P with F_M	swipe F_L with T	tap T and F_L
Enlarge	swipe P with T	swipe T with F_L	tap T and F_M	tap T and F_M
Shrink	swipe F_M with T	swipe P with T	tap T and F_M	tap T and F_M
Zoom in	swipe F_M with T	swipe F_L with T	swipe F_L with T	swipe F_L with T
Zoom out	tap T and F_L	tap T and F_L	swipe F_L with T	swipe F_L with T
Open menu	tap T and F_M	swipe F_L with T	fist	fist
Close menu	tap P with F_M	draw	fist	tap T and F_M
Save	fist + tap F_L with T	tap P with F_L	tap P with F_M	fist
Minimize	tap P with F_L	tap T and F_M	tap P with F_L	tap P with F_L
Maximize	tap and swipe F_L with T	tap P with T	tap T and F_M	tap P with F_L
Scrolling	swipe F_M with T	swipe F_L with T	swipe F_L with T	swipe F_L with T
Find	tap T and F_L	circle F_M with T	draw	tap P with F_L
Volume up	tap F_M with T	tap P with T	tap T and F_L	swipe F_M with T
Volume down	tap F_M with T	swipe F_L with T	swipe F_L with T	swipe F_L with T
Mute	tap P with T	tap T and F_L	tap F_L + draw	tap T and F_L + fist
Play	swipe F_L with T	swipe F_L with T	tap P with F_L	tap P with F_L
Pause	swipe F_L with T	tap P with F_L	tap P with F_L	tap T and F_L
Stop	tap P with T	draw	tap T and F_L	tap P with F_M
Next	swipe F_M with T	swipe F_M with T	swipe F_M with T	tap T and F_L
Previous	swipe F_M with T	swipe F_M with T	swipe F_M with T	tap T and F_L

T = Thumb; F_L = one or two fingers; F_M = three or more fingers; P = Palm

referent	p12	p13	p14
On	fist	fist	fist
Off	fist	fist	fist
Select single	tap T and F_L	tap T and F_L	tap T and F_L
Select group	tap T and F_M	tap T and F_M	tap T and F_M
Move	swipe F_M with T	tap T and F_L	swipe F_L with T
Pan	swipe F_M with T	swipe F_M with T	swipe F_M with T
Rotate	circle F_L with T	circle F_L with T	circle F_L with T
Cut	tap F_L	tap F_L	tap T and F_L
Copy	tap T and F_L	tap T and F_L	tap T and F_L
Paste	tap T and F_L	tap T and F_L	tap T and F_L
Delete	tap T and F_L	tap T and F_L	swipe F_M with T
Accept	tap T and F_L	tap T and F_L	tap T and F_L
Reject	tap P with F_M	tap P with F_M	tap T and F_L
Help	swipe F_L with T	tap P with F_L	tap T and F_M
Undo	swipe F_M with T	swipe F_M with T	swipe F_M with T
Enlarge	tap T and F_M	swipe F_L with T	swipe F_L with T
Shrink	swipe F_L with T	swipe F_L with T	swipe F_L with T
Zoom in	swipe F_L with T	swipe F_L with T	swipe F_L with T
Zoom out	swipe F_L with T	swipe F_L with T	swipe F_L with T
Open menu	tap T and F_L	tap T and F_L	tap T and F_L
Close menu	tap T and F_L	tap T and F_L	tap T and F_L
Save	swipe F_L with T	tap T and F_L	tap T and F_L
Minimize	tap P with F_L	swipe F_L with T	swipe F_L with T
Maximize	tap T and F_L	swipe F_L with T	swipe F_L with T
Scrolling	swipe F_L with T	swipe F_L with T	swipe F_L with T
Find	tap P with F_M	tap T and F_M	tap T and F_L
Volume up	swipe F_M with T	swipe F_L with T	swipe F_L with T
Volume down	tap T and F_L	tap T and F_L	swipe F_M with T
Mute	swipe F_M with T + fist	tap F_M with T	tap F_M with T
Play	tap T and F_M	tap T and F_M	tap T and F_L
Pause	tap T and F_L	tap T and F_L	tap T and F_L
Stop	tap P with F_M	fist	fist
Next	tap T and F_L	tap T and F_L	tap T and F_L
Previous	tap T and F_L	tap T and F_L	tap T and F_L

T = Thumb; F_L = one or two fingers; F_M = three or more fingers; P = Palm

referent	p15	p16
On	fist	fist
Off	fist	fist
Select single	tap T and F_L	tap T and F_L
Select group	tap T and F_M	tap T and F_M
Move	swipe F_L with T	swipe F_L with T
Pan	swipe F_M with T	swipe F_M with T
Rotate	circle F_L with T	circle F_L with T
Cut	tap T and F_L	tap P with F_M
Copy	tap T and F_L	tap T and F_L
Paste	tap T and F_L	tap T and F_L
Delete	swipe F_M with T	swipe F_M with T
Accept	tap T and F_L	tap T and F_L
Reject	tap T and F_L	tap T and F_L
Help	tap T and F_L	tap T and F_L
Undo	swipe F_M with T	swipe F_M with T
Enlarge	swipe F_L with T	swipe F_L with T
Shrink	swipe F_L with T	swipe T with F_L
Zoom in	tap T and F_L	tap T and F_L
Zoom out	swipe F_L with T	swipe F_L with T
Open menu	tap T and F_L	tap T and F_L
Close menu	tap T and F_L	tap T and F_L
Save	tap T and F_L	tap T and F_M
Minimize	swipe F_L with T	swipe F_L with T
Maximize	swipe F_L with T	swipe F_L with T
Scrolling	swipe F_L with T	swipe F_L with T
Find	tap T and F_L	tap T and F_L
Volume up	swipe F_L with T	swipe F_L with T
Volume down	swipe F_M with T	swipe F_M with T
Mute	fist	tap T and F_L
Play	tap T and F_L	tap T and F_L
Pause	tap T and F_M	tap T and F_M
Stop	fist	fist
Next	tap T and F_L	tap T and F_L
Previous	tap T and F_L	tap T and F_L

T = Thumb; F_L = one or two fingers; F_M = three or more fingers; P = Palm

Appendix B: Gesture Classification Methodology

The study had 16 participants, with each participant designing 3 gestures for each of the 34 referents. This resulted in 1,632 gestures, which were classified based on the *Descriptive Labeling* and *Chunking and Phrasing* techniques of Nielsen [35] and Buxton [2], respectively. Of the 3 gestures for each referent, users were asked to select one gesture as the preferred gesture; this left us with a third of the 1,632 gestures, equivalent to 544 gestures. Within the 544 gestures, there were 140 unique gestures, with the rest being duplicates.

While reviewing user feedback as well as user behavior from during the studies, we realized that users were often incapable of recalling the exact fingers used in the gestures they proposed. For example, they might chose a tap gesture, but forget whether the index or middle finger was tapped (with the thumb). In other cases, users made a distinction between using fewer or more fingers as a metaphor to reflect the referent, but did not care about the exact number of fingers. This led us to loosen the classification scheme, from specifying each finger used, to only differentiating between more or less fingers being used. We defined "more" to be three or more fingers, and "less" to be one or two fingers. Based on this revised classification scheme, we reduced the set to 47 unique gestures. Finally, we selected the consensus gesture for each referent, which left us with only 8 unique gestures. These 8 unique gestures accounted for 220/544 of the original set of preferred gestures. The remaining 324/544 are made up of the 39 unique gestures which did not make it into the consensus set.

With 34 referents and only 8 unique gestures, each gesture was likely to be repeated for multiple referents. To reduce the amount of conflicts, specific fingers were assigned to variations of the 8 unique gestures. We decided which fingers to use for each referent, by identifying the most commonly used fingers amongst all the gestures proposed for that referent. By creating variations this way, we ended up with a final consensus set of 16 gestures, representing 34 referents. Taking into account that the 34 referents are split into 6 categories, in separate contexts, all conflicts were resolved between gestures and referents.

Appendix C: Co-author Permission

This appendix contains the co-authors written permission to use the content of the following publications in this thesis and to have this work microfilmed:

- *User Elicitation on Single-hand Microgestures.* Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Yang, Frank Maurer. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI 2016). ACM. New York, NY, USA, 2016.

UNIVERSITY OF CALGARY

FACULTY OF SCIENCE

Department of Computer Science
2500 University Drive NW
Calgary, AB, Canada T2N 1N4
ucalgary.ca

March 24th, 2017

I, Teddy Seyed, grant Edwin Chan full permission to use the content of the following co-authored publications in his M.Sc. dissertation and to have this work microfilmed.

- *User Elicitation on Single-hand Microgestures*. Edwin Chan, Xing-Dong Yang, Teddy Seyed, Wolfgang Stuerzlinger, Frank Maurer. In Proceedings of the ACM Conference on Human Factors in Computing Systems 2016 (CHI '16), San Jose, USA, 2016.

UNIVERSITY OF CALGARY

FACULTY OF SCIENCE

Department of Computer Science
2500 University Drive NW
Calgary, AB, Canada T2N 1N4
ucalgary.ca

March 24th, 2017

I, Xing-Dong Yang, grant Edwin Chan full permission to use the content of the following co-authored publications in his M.Sc. dissertation and to have this work microfilmed.

- *User Elicitation on Single-hand Microgestures*. Edwin Chan, Xing-Dong Yang, Teddy Seyed, Wolfgang Stuerzlinger, Frank Maurer. In Proceedings of the ACM Conference on Human Factors in Computing Systems 2016 (CHI '16), San Jose, USA, 2016.

FACULTY OF SCIENCE

Department of Computer Science
2500 University Drive NW
Calgary, AB, Canada T2N 1N4
ucalgary.ca

March 24th, 2017

I, ____Wolfgang Sturzlinger____, grant Edwin Chan full permission to use the content of the following co-authored publications in his M.Sc. dissertation and to have this work microfilmed.

- *User Elicitation on Single-hand Microgestures*. Edwin Chan, Xing-Dong Yang, Teddy Seyed, Wolfgang Stuerzlinger, Frank Maurer. In Proceedings of the ACM Conference on Human Factors in Computing Systems 2016 (CHI '16), San Jose, USA, 2016.