A Taxonomy of Freehand Grasping Patterns in Virtual Reality

Andreea Dalia Blaga

BIRMINGHAM CITY
University

DMT Lab

School of Computing and Digital Technology

Faculty of Computing, Engineering and the Built Environment

Birmingham City University

A thesis presented for the degree of

Doctor of Philosophy

Acknowledgements

This thesis would not have been possible without the support of many people.

To my supervisors, Prof. Ian Williams, Dr Maite Frutos-Pascual and Prof. Chris Creed, I cannot thank you enough for your guidance and support. Special thanks to Ian for being a great mentor since my first university years, taking me on as a PhD student and having faith in me over the years, always willing to offer advice on both personal and professional levels. Maite, thank you for always believing in me, guiding me and making stressful periods easier. Thanks to all members of the DMT Lab who inspired and helped me complete this work. Thank you Rehan Bhana for always encouraging me to keep going and being an amazing mentor.

To my parents and family, thank you for teaching me to always follow my dreams, supporting and motivating me to achieve my goals. To Iulia and Dan Lazar, thank you for your overwhelming generosity, kindness and for taking me in as one of your own. To my best friends, Ingrid, Irina and Andrada, thank you for always being there for me, for the good and bad times, celebrating every small victory with me.

Last but definitely not least, my partner, Bogdan Lazar for always being there for me. I can't thank you enough for making me laugh in stressful times, always listening and ready to help in any way. Thank you for all the sacrifices you made during this time. Your love and support made this journey a lot easier.

Abstract

Grasping is the most natural and primary interaction paradigm people perform every day, which allows us to pick up and manipulate objects around us such as drinking a cup of coffee or writing with a pen. Grasping has been highly explored in real environments, to understand and structure the way people grasp and interact with objects by presenting categories, models and theories for grasping approach. Due to the complexity of the human hand, classifying grasping knowledge to provide meaningful insights is a challenging task, which led to researchers developing grasp taxonomies to provide guidelines for emerging grasping work (such as in anthropology, robotics and hand surgery) in a systematic way.

While this body of work exists for real grasping, the nuances of grasping transfer in virtual environments is unexplored. The emerging development of robust hand tracking sensors for virtual devices now allow the development of grasp models that enable VR to simulate real grasping interactions. However, present work has not yet explored the differences and nuances that are present in virtual grasping compared to real object grasping, which means that virtual systems that create grasping models based on real grasping knowledge, might make assumptions which are yet to be proven true or untrue around the way users intuitively grasp and interact with virtual objects.

To address this, this thesis presents the first user elicitation studies to explore grasping patterns directly in VR. The first study presents main similarities and differences between real and virtual object grasping, the second study furthers this by exploring how virtual object shape influences grasping patterns, the third study focuses on visual thermal cues and how this influences grasp metrics, and the fourth study focuses on understanding other object characteristics such as stability and complexity and how they influence grasps in VR. To provide structured insights on grasping interactions in VR, the results are synthesized in the first VR Taxonomy of Grasp Types, developed following current methods for developing grasping and HCI taxonomies and re-iterated to present an updated and more complete taxonomy.

Results show that users appear to mimic real grasping behaviour in VR, however they also illustrate that users present issues around object size estimation and generally a lower variability in grasp types is used. The taxonomy shows that only five grasps account for the majority of grasp data in VR, which can be used for computer systems aiming to achieve natural and intuitive interactions at lower computational cost. Further, findings show that virtual object characteristics such as shape, stability and complexity as well as visual cues for temperature influence grasp metrics such as aperture, category, type, location and dimension. These changes in grasping patterns together with virtual object categorisation methods can be used to inform design decisions when developing intuitive interactions and virtual objects and environments and therefore taking a step forward in achieving natural grasping interaction in VR.

Table of Contents

List of Figures

List of Tables

1 | Introduction

1.1 Motivation

Virtual Reality (VR) is in a period of strong growth, with the number of virtual environments increasing day by day. One significant factor that is fuelling this advancement is the unprecedented growth in consumer availability and use, with companies such as Meta, Microsoft and Samsung enabling accessible VR experiences for the masses. VR applications are now increasingly used in the entertainment and gaming world which include VR social platforms (Facebook Spaces), immersive cinemas (IMAX), museum tours (British Museum), live concerts, live sports games (Meta's Oculus venues) and 3D immersive games that replicate traditional game genres in VR.

This growth in consumer available VR hardware and software attracted the research community to investigate the use of VR outside entertainment applications, to support decision making and enable innovation, while taking workforce training to the next level by enabling highly immersive environments (Frutos-Pascual, Harrison, Creed, & Williams, 2019). The high immersion levels and the ability to replicate real scenarios in VR has allowed the development of VR training applications and simulations (Figure 1.1), which bring several advantages when compared to traditional training and learning, including the ability to simulate any situation without exposing trainees to its risks and the ability to repeat a training session for an unlimited number of times (Ragan et al., 2015). Hence, from flight simulations (X. Sun, Liu, Tian, Wu, & Gao, 2020), surgical training techniques (Nayer, Murdock, Dharia, & Belyea, 2020), psychological therapies (Opris et al., 2012) or fire evacuation simulations (Lawson, Roper, Shaw, Hsieh, & Cobb, 2020), the possibilities are rapidly developing.

Figure 1.1: Controller-based surgical training platform from Osso VR (*Osso VR Virtual Reality Surgical Training Platform*, n.d.).

Yet the usability and effectiveness of these simulations is highly dependent on several factors, which have shown to play a key role in mimicking real scenarios and significantly impact user quality of experience (Hudson, Matson-Barkat, Pallamin, & Jegou, 2018). Immersion is known as the perception of being physically present in a non-physical world, and is achieved by surrounding the user by visual, auditory and other stimuli to "block out the physical world" (Biocca, 1992). Beyond immersion, the main component of VR is interaction (Heim, 2000), which allows the user to interact with the virtual objects and improves presence in these environments (Hudson, Matson-Barkat, Pallamin, & Jegou, 2019), which is particularly important for virtual environments which aim to mimic real scenarios for training and simulations. In real environments, people are accustomed to interactions between people and surrounding objects where they receive information via multiple sense organs in ways of seeing, listening, speaking, touching and tasking. This inspired the development of various interaction methods in VR, researchers focusing on providing real-time interactivity through speech, head movements, gaze, touch or 3D hand interactions (Figure 1.2).

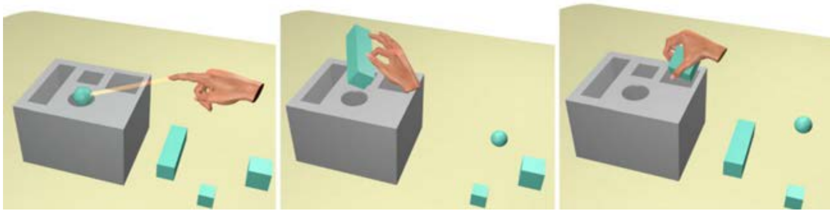


Figure 1.2: Virtual hand-object interactions proposed by (Tian et al., 2019).

Hand interaction has gained popularity in VR with the rapid technological advancements that allow users to interact with virtual systems using their hands, due to human hand dexterity and humans' ability to use their hands for acquiring and manipulating objects with ease. Commonly, hand-held controllers are used as the standard interaction method for VR interactions, especially for consumer-available and entertainment VR. However, controllers have shown to be limited in providing natural and intuitive interactions, users often reporting that interactions are not intuitive and require a longer learning curve (Tanjung, Farhan, Siregar, Panjaitan, & Fahmi, 2020), which is particularly important for VR scenarios where mimicking reality is important for knowledge transfer.

Suppose you work in the manufacturing industry, training for a challenging assembly task. Instead of reading procedures or watching others perform the task, VR now allows the development of a training scenario where you can practice the important steps multiple times, allowing you to learn without risking your and your co-worker's safety. High immersion levels through visual and auditory stimuli help you feel present in this environment and provide feedback for when you make a mistake. The aim of this training is for you to translate the assembly skills learned during this experience to a technique you will confidently perform in the real scenario. Now to achieve this, you would like to be able to interact with the virtual environment in the most natural and intuitive manner, which would not require you to learn new interactions but rather allow you to focus on the goal of the training session. Naturally, the most comfortable interaction technique would be to use your hands in the same way you use them in real environments. This would provide a learning experience where the consequences are not real but with the benefits of hands-on learning, which has been studied in psychology and education and showed numerous learning benefits such as increased motivation, improved on-the-job performance and shorter learning curves (Cridlin, 2007).

Hand interactions that take advantage of the dexterous versatility of the human hand have been highly explored within the Human-Computer Interaction (HCI) community (Q. Wang, Kang, & Kristensson, 2021) initially through the use of wearables such as gloves(Maldonado & Zetzsche, 2021), however previous work showed that bare hand interactions (e.g. without using devices to augment the hand) mitigate some of the limitations associated with wearables such as fatigue and discomfort and therefore have been linked to ease of access and naturalness (Oudah, Al-Naji, & Chahl, 2020). When creating new bare hand interactive systems, several studies rely on predefined gestures, which are generally designed for optimal recognition rather than naturalness, being often arbitrary and not intuitive enough (Piumsomboon, Clark, Billinghurst, & Cockburn, 2013), which led researchers to focus on physical interaction paradigms for VR environments where natural and intuitive interaction is required.

Grasping is the primary and most frequent physical interaction technique people perform in everyday life and is defined as every static posture at which an object can be held securely with a single hand. Virtual grasping has extensively been explored as a technical and computational challenge, however with current approaches, users are often trained to use particular grasps, with the design considerations and grasping constraints used in these solutions being applied from the body of knowledge available in real object grasping. This approach assumes that in order to achieve intuitive and natural grasping for training and simulations, virtual grasp models need to replicate real grasping movements. In real environments, the hand pose during grasping is influenced by both visual perception and haptic feedback that inform us on the shape, weight, texture or temperature of a real object, which we then use to make a decision on how to perform a grasp that ensures stability for the intended task. However, current VR technology is still limited in offering haptic feedback (Islam & Lim, 2022), with the majority of grasping interaction decisions being made based on visual perception only. Now this introduces the question of whether the limitations we are currently facing in VR influence intuitive grasping patterns, and whether or not virtual grasping models should completely mimic real grasping patterns to achieve natural and intuitive interactions in VR.

Evaluating grasping patterns directly in VR will aid in answering this question and understanding the intuitive hand poses users perform in VR when haptic feedback is missing as well as how these grasping patterns change in VR for common factors that influence user grasping interaction in real environments such as object characteristics, task and visual thermal cues. Nonetheless, the complexity and variety of uses of the human hand makes the categorisation and classification of hand function a challenging task, still, synthesising grasps in taxonomies has shown to be beneficial for defining common terminology and informing new research directions in real grasping research. Developing a VR grasp taxonomy would therefore inform the design of virtual grasping models and more natural and intuitive VR environments and objects, as well as providing an overview of key user behaviours, limitations and problems when grasping in VR, taking a step forward in achieving natural and intuitive interactions in VR, which could also contribute to current research trends that aim to move the metaverse from science fiction to an upcoming reality (Y. Wang et al., 2022).

1.2 Aim and Objectives

The aim of this work is to evaluate grasping patterns in VR and develop the first VR Taxonomy of Grasp Types. This is achieved through the following objectives:

1. Review and determine current trends in 3D hand interaction and real grasping research.
2. Define a methodology for collecting grasping patterns in VR suitable for determining grasping trends and taxonomies.
3. Explore and quantify the differences and similarities between grasping real objects and grasping virtual objects.
4. Measure the impact of object characteristics and tasks on grasping metrics in VR.
5. Evaluate differences in grasping approach based on visual cues for avatar and thermal feedback representation.
6. Synthesize grasp instances in the first VR Taxonomy of Grasp Types.
7. Define and synthesize grasp patterns and potential applications of the taxonomy for virtual environment object grasping work.

1.3 Thesis Structure

The aim of this work is to evaluate grasping patterns in VR and develop the first VR Taxonomy of Grasp Types. Firstly, current methods for developing real grasping and HCI taxonomies as well as grasp metrics for evaluating hand pose in real environments were reviewed. Based on these and following adaptations to mitigate VR limitations, a novel method for developing the first VR Taxonomy of

Grasp Types was proposed. Four user studies are then conducted to collect grasp data under various conditions and synthesize the results in VR grasp taxonomies. Findings from user studies are employed to inform grasping interaction design decisions for achieving more intuitive and natural interactions in VR. The objectives of this thesis are achieved throughout nine chapters (see Figure 1.3) as follows:

In **Chapter 2**, the background research into 3D hand interaction in virtual environments is presented. Firstly, an overview of input devices used for hand interaction is presented, describing controller-based interactions and hand tracking based interactions, where wearable-based and freehand interactions are detailed. This is followed by an overview of interaction paradigms for freehand interactions, discussing advantages and limitations of gesture-based interaction and physical interaction. Finally, current methods for virtual grasping are discussed, together with problems and limitations, presenting the need for a more systematic exploration of grasping in VR.

In **Chapter 3**, the background research into existing real object grasping taxonomies is presented. It first presents the definition of grasping. Then, the biomechanics of the hand during grasping in real environments are presented, followed by measures used to analyse grasp poses for real objects. This is followed by a detailed overview of grasp taxonomies where the types of classification are presented. Finally, the most up to date real grasp taxonomy in literature, which is used for classifying grasps in this thesis, is presented in detail, together with terminology and use.

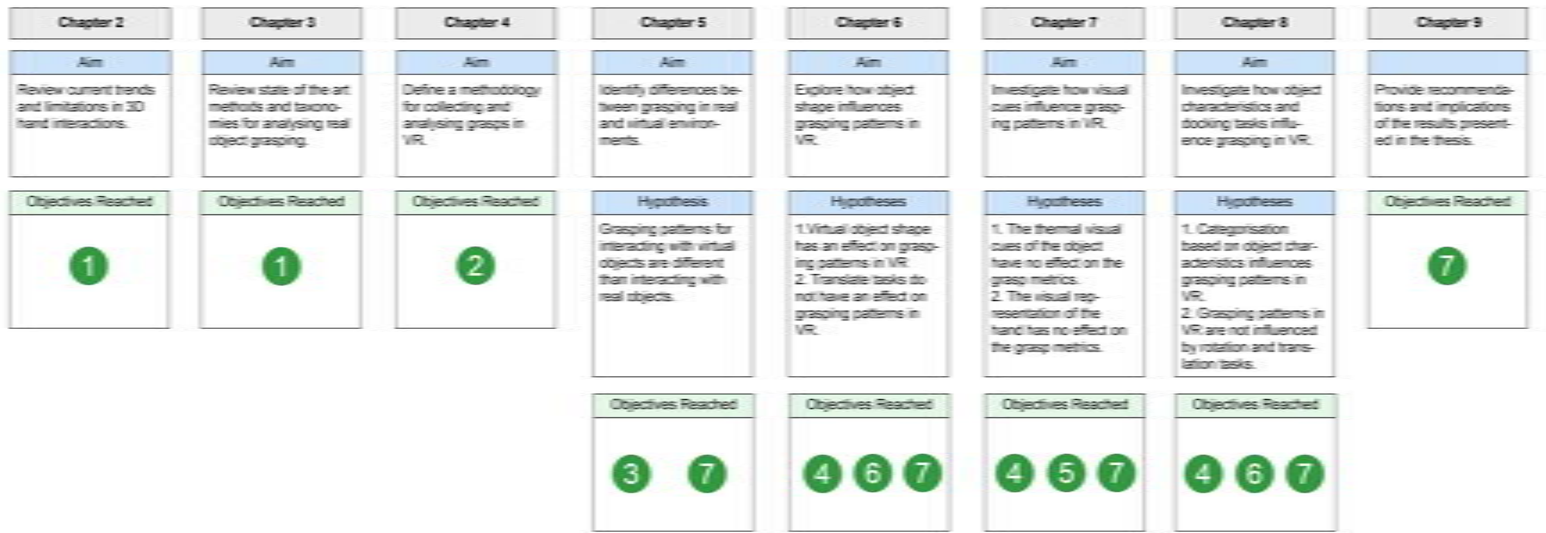


Figure 1.3: Overview of the aim of each chapter presented in this thesis, together with hypotheses of user studies. For each chapter, objectives reached is presented, which shows which of the overall thesis objectives were achieved in each chapter. The numbers in this figure connect to the numbers of each objective presented in Section 1.2.

In **Chapter 4**, the current methods for developing HCI taxonomies are reviewed and a novel method for developing a VR taxonomy of grasps is proposed based on the real object taxonomy literature in Chapter 3. First, it provides an overview of taxonomies in HCI, then it describes data collection methods used for developing taxonomies. This is then followed by an overview of the proposed method, which is based on the reviewed literature. Next, the baseline environment for the user studies presented in this thesis, together with the grasp metrics and labelling methodologies are detailed. Finally, the modifications to the fundamental method for each user experiment in Chapters 5-8 are presented.

In **Chapter 5**, a first user experiment to explore differences between grasping metrics in real and virtual environments is presented. Time to grasp, grasp aperture and grasp labels are reported for both real and virtual objects. Key similarities and differences in grasp metrics are discussed.

In **Chapter 6**, grasping patterns are explored for different object shapes and simple translate tasks, following assumptions made in Chapter 5 and real grasping literature, namely that virtual object shape, might influence grasping patterns in VR. Grasp aperture and grasp labels are collected and analysed for each object shape, with results being synthesized in the first VR Taxonomy of Grasp Types.

In **Chapter 7**, a first user study to analyse grasping patterns for visual cues representing thermal haptic feedback and user hand avatar is presented. Assumptions from Chapters 5 and 6 are addressed and grasp location, grasp aperture and grasp labels are analysed for different visual thermal cues to understand how thermal haptic feedback and realism of hand avatar influences grasping approach in VR.

In **Chapter 8**, a first user study to analyse the effect of different categorisation methods which explore not only virtual object shape but virtual object stability and complexity is presented, inspired by findings in Chapters 5-7 showing that virtual object characteristics influence grasping approach in VR. To further the work in Chapter 6, the grasp patterns are analysed during a mixed docking task (rotation and translation) to evaluate how grasping patterns change for different tasks, inspired by findings in real grasping literature. The results are synthesized in an updated, more complete VR Taxonomy of Grasp Types, complementary to the taxonomy presented in Chapter 6.

In **Chapter 9**, a set of recommendations based on findings from Chapters 5-8 is presented. Findings in this work are discussed together with wider implications for the VR and HCI community. Finally, limitations and future work is presented.

1.4 Contributions

The primary contribution of this thesis is the first evaluation of freehand grasping patterns in VR for common influencing factors, synthesised in the first VR Taxonomy of Grasp Types. In achieving this, a number of other contributions are made:

- Novel methodology for evaluating freehand grasping in VR (Chapter 4)
- First study to analyse differences in grasp metrics between real and virtual environments (Chapter 5 and (Blaga, Frutos-Pascual, Creed, & Williams, 2021b))
- Development of the first VR Taxonomy of Grasp Types through synthesising the results from a comprehensive analysis of grasping patterns for object shape (Chapter 6 and submitted and under review to IJHCI)
- Analysis of grasping patterns in VR based on visual cues for hand representation and thermal haptic feedback (Chapter 7 and (Blaga, Frutos-Pascual, Creed, & Williams, 2020))

- Novel virtual categorisation methods and development of an updated, more complete VR Taxonomy of Grasp Types to reflect changes in grasp metrics based on virtual object characteristics (Chapter 8 and (Blaga, Frutos-Pascual, Creed, & Williams, 2021a, 2021c)

1.5 Published Papers

The following papers have been published as part of this work:

- Andreea Dalia Blaga, Maite Frutos-Pascual, Chris Creed, and Ian Williams. 2020. Too Hot to Handle: An Evaluation of the Effect of Thermal Visual Representation on User Grasping Interaction in Virtual Reality. 2020 CHI Conference on Human Factors in Computing Systems (CHI ’20). **[Core A* Ranking]**
- Andreea Dalia Blaga, Maite Frutos-Pascual, Chris Creed and Ian Williams. 2021. Freehand Grasping: An Analysis of Grasping for Docking Tasks in Virtual Reality. 2021 IEEE Virtual Reality and 3D User Interfaces (VR), 2021 **[Core A Ranking]**
- Andreea Dalia Blaga, Maite Frutos-Pascual, Chris Creed and Ian Williams. 2021. A Grasp on Reality: Understanding Grasping Patterns for Object Interaction in Real and Virtual Environments. 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), 2021 **[Core A* Ranking]**
- Andreea Dalia Blaga, Maite Frutos-Pascual, Chris Creed, and Ian Williams. 2021. Virtual Object Categorisation Methods: Towards a Richer Understanding of Object Grasping for Virtual Reality. 27th ACM Symposium on Virtual Reality Software and Technology (VRST ’21) **[Core A Ranking]**

1.6 COVID-19

This work was undertaken before and during the period of the COVID-19 pandemic. User studies and data collected during and after the pandemic were following COVID-19 safety guidelines.

2 | Hand Interaction in VR

2.1 Introduction

While immersion has shown to play a key role in achieving realistic experiences in VR (Tan, Niu, & Zhang, 2020), another important aspect that influences presence and realism in the VR experience is interaction. Interaction in VR is often described as the ability of the user to move within the virtual world and to interact with the objects of the virtual world (Bostan, 2006). People are accustomed to interactions between people and surrounding objects in daily life where they receive information via multiple sense organs in ways of seeing, listening, speaking, touching and tasking (Shen, 2021). This inspired the development of various interaction methods in VR, with researchers focusing on providing real-time interactivity that allows the user to interact with a computer interface in a similar way that they interact in real environments, to allow high immersion and presence in VR (Khenak, Vézien, & Bourdot, 2020). These interaction methods make use of speech (Azizo, Mohamed, Siang, & Isham, 2020), head movements (Yu, Liang, Zhang, & Xu, 2019), eye-gazing (Piumsomboon, Lee, Lindeman, & Billinghurst, 2017), touch (Y. R. Kim, Choi, Chang, & Kim, 2020) and hands.

Hand interaction has gained popularity in VR with the rapid technological advancements that allow users to interact with virtual systems using their hands, due to human and dexterity and humans' ability to use their hands for acquiring and manipulating objects with ease (Vogel & Balakrishnan, 2005). Thus, for VR systems that aim to replicate real scenarios in VR, researchers focused on using 3D hand interaction tools such as hand-held controllers and hand tracking sensors to allow interactions that are easy to learn in virtual environments. Taking into consideration the dexterous versatility of the human hand, researchers have explored different interaction paradigms to accommodate the needs and limitations of virtual systems, with the main paradigms used in VR today being gestures and physical interaction.

This section presents an overview of input devices used for 3D hand interaction as well as interaction paradigms for state-of-the-art VR. Section 2.2 presents input devices for hand interaction with the main categories being controller-based interaction and hand tracking interaction (wearable-based interaction and freehand interaction). Section 2.3 presents freehand interaction paradigms, detailing benefits and limitations of gesture-based interaction and physical interaction. A more detailed overview of current trends in physical interaction, namely virtual grasping is presented in Section 2.4.

2.2 Input Devices for Hand Interaction

Hand interaction with virtual objects can be achieved using a variety of techniques in VR. Since standard devices such as keyboard and mouse are difficult to use in a highly immersive VR environment (Jayaram, Vance, Gadh, Jayaram, & Srinivasan, 2001), researchers focused on creating alternative devices that retain the authentic and universal sense of reality by preserving close ergonomic similarities with human physical and manual dexterity and agility (Carmeli, Patish, & Coleman, 2003). These alternative devices are now the state-of-the-art in immersive VR interactions and can be divided in two main categories: controller-based interactions and hand tracking interactions. The following sections present an overview of these types of input devices together with their use and limitations.

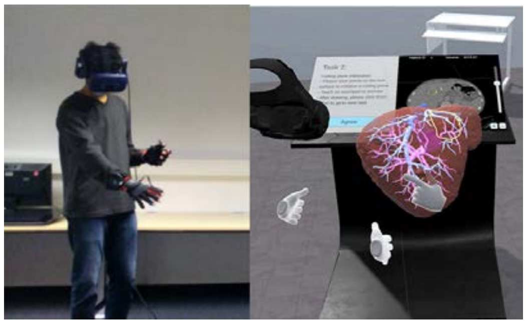
2.2.1 Controller-based Interaction

The most popular input devices for hand interaction in fully immersive environments are controllers, due to their accuracy and low cost, being mainly used to improve the sense of immersion through interaction in virtual environments (Choi, Ofek, Benko, Sinclair, & Holz, 2018). Controllers may be wired or wireless, are hand worn and provide discrete input in the form of buttons and continuous input by top-mounted joysticks or touch-pads which provide tracking of the position and orientation of users' hands with high accuracy and fast recognition speed (Zhang et al., 2018), which has shown to increase user presence during interaction in VR (Caggianese, Gallo, & Neroni, 2019; Tanjung et al., 2020).

Due to these interaction opportunities proposed by controllers, top VR HMD companies such as Meta, HTC and PlayStation introduced controllers as the main interaction tools to accompany their headsets, making them easily available and affordable at consumer level. This led to an increased use of controllers for VR entertainment applications (Vogel, Lubos, & Steinicke, 2018; H. Park, Faghihi, Dixit, Vaid, & McNamara, 2021). This increased popularity of controllers was also evident in the HCI community, where researchers focused on developing and evaluating controller-based interactions to achieve highly interactive systems (Suznjevic, Mandurov, & Matijasevic, 2017) by either mimicking interactions from existing UIs in VR (mouse interactions such as pointing and selecting (Capece, Erra, & Grippa, 2018)) or mimicking hand interaction behaviour from real environments (picking and manipulating a virtual object (Suznjevic et al., 2017)).

However, researchers found that controller-based interactions might present challenges for users when learning how to correctly hold and manipulate the controller for specific interactions (Tanjung et al., 2020). For example, Hartney et al. (Hartney et al., 2019) developed an interactive application for upper-limb train-

(a) Controller-based Interaction



(b) Wearable-based Interaction

Figure 2.1: Controller-based interaction for a virtual scenario mimicking real interactions as presented in (Hartney et al., 2019) and wearable-based interaction for a virtual scenario mimicking real interactions presented in (Chheang et al., 2021)

ing in injured patients, where users were asked to perform daily tasks, such as cleaning a virtual window using controllers (Figure 2.1 a). They showed that even though users were educated on how to use the controllers, users reported that the interaction took longer to learn and was challenging, due to the movements required being very different from real-life activities, which was also found in (D. Chen, Liu, & Ren, 2018; Lougiakis, Katifori, Roussou, & Ioannidis, 2020). Moreover, researchers showed that the design between the most popular commercially available controllers is incongruent, leading to inconsistent levels of accessibility which hinder intuitive interactions in VR (Cook, Dissanayake, & Kaur, 2019).

While controllers are currently the most common input device for immersive hand interactions, they have a higher correlation with individual bespoke functionalities than with a standardised relationship for authentic HCI. These findings are particularly important to consider for virtual environments where replicating real tasks and behaviour is needed, such as for training and simulations (Gonzalez & Garnique, 2018), where input devices that do not rely on buttons and touchscreens and in turn propose more natural and intuitive approaches can be considered.

2.2.2 Hand Tracking Interaction

Another important branch of input devices for immersive interaction is hand tracking. Hand tracking focuses on allowing users to perform movements similar to interactions in real environments, by tracking hand movements such as position and orientation of the palm and fingers in 3D. This interaction technique allows direct interactions between the hand and virtual objects and thus gained popularity for VR systems that mimic realistic scenarios such as training (Levin, Magdalon, Michaelsen, & Quevedo, 2015) and simulations (Q. Wang et al., 2021). The approaches for allowing hand tracking interactions in VR are diverse, however they can be split in two main categories: wearable-based interaction, where wearable devices are placed on the hand/or arm and freehand interaction where the hand is not augmented with additional sensory or feedback devices. The next sections present an overview of these two types of hand tracking interactions, presenting their use and limitations.

2.2.2.1 Wearable-based Interaction

Wearable-based interactions in VR utilise wearable sensors or tracking markers placed on the hand or arm used for recording data related to user hand configuration and motion (such as the bending angle and level of adduction of each finger) (Dipietro, Sabatini, & Dario, 2008). The most common wearable devices used in VR are data gloves, which gained popularity due to allowing interaction paradigms where hand muscles are engaged in a similar way as humans use them for everyday tasks, which cannot be achieved with traditional VR controllers (Maldonado & Zetzsche, 2021) (Figure 2.1 b). This is evident in the work of Almeida et al. (Almeida et al., 2019) who compared a data glove to a controller for virtual object interaction and showed that the sense of embodiment and speed of completion were significantly higher with the data glove, which was also found in (J. Lee, Sinclair, Gonzalez-Franco, Ofek, & Holz, 2019). Moreover, users often reported that the glove interaction was more intuitive as it allowed them to replicate movements they are familiar with from real interactions.

Due to the nature of the interaction paradigms proposed by wearable devices, researchers focused on using them for training of the hand in injured patients (Tsoupikova et al., 2014) and showed significant results for improving hand movements for daily tasks as well as improving muscle balance and functional parameters (Reyes-Guzman et al., 2015). This potential of wearable devices to provide interaction techniques that are easy to learn and intuitive, led to an increased popularity of wearable input for various VR applications such as medical rehabilitation (Levin et al., 2015), simulations (Moehring & Froehlich, 2011), training (Cao, Gao, Wang, & Li, 2016), collaborative VR (Chheang et al., 2021), robotics (Fu, Fu, Guo, Guo, & Li, 2020), sign language understanding (Anupama, Usha, Madhushankar, Vivek, & Kulkarni, 2021), entertainment (Adamo-Villani & Wilbur, 2007) and mental health therapy (Q. Wang et al., 2021).

However, while using data gloves have shown to be beneficial for applications where engaging the muscles of the hand is important for creating a realistic experience, data gloves that provide high accuracy are usually expensive and therefore not easily accessible for consumer use (Han, 2010). This has led researchers to develop glove systems using low-cost sensors for applications that are created to be widely used by the population (Cao et al., 2016), however researchers often find imperfections in hand tracking which affect the interaction quality (Borst & Indugula, 2005). To reduce these inaccuracies, most wearable devices need to be calibrated for particular users, which has shown to be a time-consuming and tedious process (Dipietro et al., 2008; Levin, Magdalon, Michaelsen, & Quevedo, 2008). Another important limitation to consider for realistic VR environments

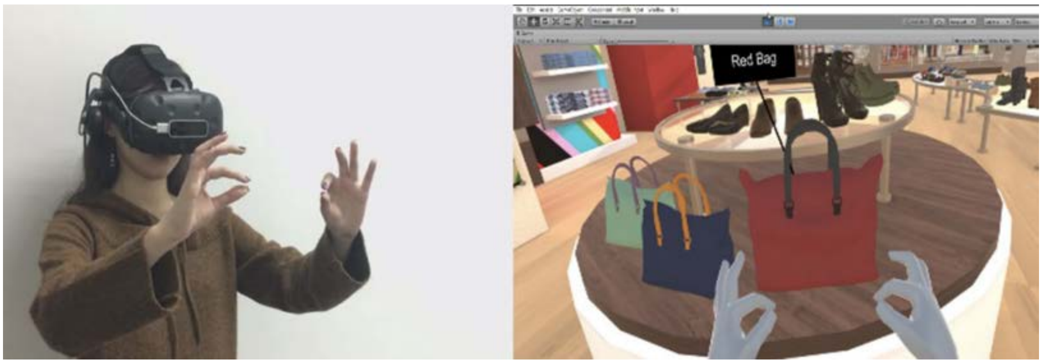


Figure 2.2: Freehand gesture-based interaction for a VR shopping application (H. Wu et al., 2019)

is that the weight of wearable devices on the hand has shown to cause fatigue during interactions and break the presence in immersive scenarios (Levin et al., 2008). To address this, researchers considered tracking hand movements without augmenting the hands, also known as freehand interaction.

2.2.2.2 Freehand Interaction

Freehand interactions, also known as bare hand interactions, have been explored as an alternative to wearable-based interaction (Figure 2.2) to allow manipulation of virtual objects without augmenting the hand with wearables or controllers and are the currently most explored interactions with head worn immersive displays (Spittle, Frutos-Pascual, Creed, & Williams, 2022). Generally, the term freehand is defined in Oxford Dictionary as:

Definition 2.2.1 (Freehand) *Drawn or executed by hand without guiding instruments, measurements, or other aids.*

This type of interaction is achieved using motion capture systems, which record and digitize the position and orientation of the hand or fingers using cameras and computer vision algorithms (Oudah et al., 2020) and depth (Soh, Choi, Park, &

Yang, 2013) and infrared sensors (Pfeuffer, Mayer, Mardanbegi, & Gellersen, 2017). The recorded movement can then be used directly to animate a virtual avatar of the user allowing them to see their hands in the virtual environment, which increases the sense of presence in VR (Coburn, Freeman, & Salmon, 2017).

This ability to interact with a virtual system without touching anything has created opportunities for use in medical applications, reducing the risk of contamination and time of surgery by allowing medical staff to interact with medical images by only moving their hand to perform some gestures (Ameur, Ben Khalifa, & Bouhlel, 2020). In addition, these interactions have shown to be very intuitive in applications that require direct manipulations of 3D objects (M. Kim & Lee, 2016), due to closely mimicking real world interactions such as moving our hands instead of pressing buttons on a controller (Tung et al., 2015) and allowing a shorter learning curve (Mu & Sourin, 2021). When compared to wearable-based interactions, users showed higher engagement and a stronger preference for free-hand interactions, as well as high immersion levels (Lages, Nabiyouni, & Arantes, 2016) and improved overall usability of virtual systems (M. Kim & Lee, 2016). The above mentioned also increased freehand interactions' popularity within the HCI community, researchers using this type of interaction for VR environments where intuitive interaction is important such as for assembly tasks (Mu & Sourin, 2021), crime scene investigation simulation (Datcu & Lukosch, 2013) or realistic games (Voigt-Antons, Kojic, Ali, & Moeller, 2020).

While freehand interactions have shown to provide benefits for allowing intuitive interactions in VR, technical limitations of infrared sensors have been associated with latency issues that might affect interaction accuracy (Silva, Abreu, de Almeida, Teichrieb, & Ramalho, 2013), which can introduce difficulties in interaction with virtual objects (M. Kim & Lee, 2016). However current devices have shown improvements in interaction accuracy (Guzsvinecz, Szucs, & Sik-Lanyi, 2019), making freehand interactions appropriate for fine bare hand manipulations in VR (Hameed, Khan, Kumar, Arain, & Hassan, 2017).

2.3 Freehand Interaction Paradigms

To allow intuitive freehand interactions that are customisable based on the interaction needed in each virtual scenario, researchers focused on complementing hand tracking with different interaction paradigms. Interaction paradigms are models or patterns of HCI in which the unconstrained pose of the users' real hand is mapped to an action in the virtual environment. Two major families of interaction paradigms are currently used in VR: gesture-based interactions and physics-based interactions, which will be described in the following sections.

2.3.1 Gesture-based Interaction

Gesture-based interactions are defined as a posture or movement of the user's upper limbs, through which people usually express interaction intentions and send out corresponding interactive information (Y. Li, Huang, Tian, Wang, & Dai, 2019). This type of interaction is common in a variety of application domains such as sign language recognition (Anjo, Pizzolato, & Feuerstack, 2012), training (Soh et al., 2013), simulations (Datcu & Lukosch, 2013), assembly (Mu & Sourin, 2021) and entertainment (Lages et al., 2016). One important benefit of using gestures for freehand interaction is that gestures can be customised based on the requirements of the application. For example, Cauchard et al. (Cauchard et al., 2019) integrated gesture-based interaction in a UI for controlling a drone, and only developed two simple gestures relevant for the goal of the application: hover (a flat open hand for pointing and dwelling) and push (the hand rotating left to right for navigating through the UI menu). Accordingly, other works explored gesture-based interactions for 3D selection, translation, rotation (Soh et al., 2013), pointing (Lages et al., 2016) or resizing of virtual objects (Piumsomboon et al., 2013). While gesture-based interactions have shown to maximise interaction enjoyment (Waldow, Misiak, Derichs, Clausen, & Fuhrmann, 2018), when defining the required interactions, researchers predominantly focused on optimal recognition rather than naturalness, which has shown to create challenges in VR interaction.

One of the most important challenges is the need for training users to learn the gestures before interacting with the virtual environment (Cauchard et al., 2019). These training sessions require the user to repeat the same gesture multiple times to understand the link between hand movements and the action performed in VR, which have shown to result in fatigue and discomfort due to the unnaturalness of the movements performed by the hand (Franco & Cabral, 2019; Xu, Liang, He, & Wang, 2019).

However, even when users are trained to perform specific hand gestures in VR, participants reported confusion regarding ways to interact in the virtual environments, which also leads to low accuracy in interaction (Soh et al., 2013), and breaks the immersion and user engagement in VR (Lages et al., 2016). Better results were achieved with users with previous experience in virtual systems or when multiple sessions of training were implemented (Datcu & Lukosch, 2013), showing that gesture-based interactions require long adaptation periods and should be used with caution in virtual systems developed for users without experience in virtual systems or where long training and learning periods should be avoided.

Another important challenge of gesture-based interaction is handling the ambiguous resulting effect of the interactions by designing gestures that clearly relate to a specific action that users can remember and are familiar with. For example, in real life, a single gesture can be used to perform multiple tasks, however, in VR, if the same gesture activates different operations on a virtual object, this leads to operation ambiguity and user confusion (D. L. Chen, Balakrishnan, & Grossman, 2020). This was found in the work of Arora et al. (Arora, Kazi, Kaufman, Li, & Singh, 2019) who focused on designing gestures based on user preference and found that users proposed the same gesture for different tasks. One solution emphasized by Ghomi et al. (Ghomi, Huot, Bau, Beaudouin-Lafon, & Mackay, 2013) is to use a pre-defined mapping that assigns each gestural manipulation to a unique operation. However, this method emphasizes the limitations discussed above, as users are forced to remember the gestural commands and the vast possibilities of operations makes the method unscalable. While gesture metaphors have proven useful for specific VR applications (Tung et al., 2015), for VR scenarios which aim to replicate user behaviour from real environments to allow higher immersion levels in VR (Covarrubias & Bordegoni, 2015), researchers focused on physical interaction paradigms.

2.3.2 Physical Interaction

Physical interactions are defined as the interactions that act physically on the virtual object as if it was a real object, for example pushing, pulling, lifting and grasping.

When compared to gesture-based interactions, this interaction paradigm has shown to be more realistic, providing higher levels of presence in VR (Voigt-Antons et al., 2020). For example, in a study comparing intuitive grasping of virtual objects to interacting with a virtual UI for typing, researchers found that during the grasping tasks participants experienced higher realism and presence and lower arousal as compared to the typing task (Cameron et al., 2011).

Participants reported lower mental demand during the grasping task showing that physical interactions are easier to perform due to replicating hand movements users are familiar with from real environments, and thus facilitate immersion in VR. A similar result was found by Zhang et al. (Zhang et al., 2018) who compared gesture-based interaction to physical interactions in virtual environments and found that physical interactions lead to a better performance in tasks than gestural interactions. Therefore, for virtual environments where immersion, realism and performance are highly important, researchers focused on developing physical interactions to allow users to grasp and manipulate virtual objects in an intuitive way (Prachyabrued & Borst, 2014). However, state-of-the-art grasping algorithms are still limited and have not yet succeeded in achieving natural and intuitive grasping in VR (Verschoor, Lobo, & Otaduy, 2018). The next section presents current trends in virtual grasping together with their limitations and challenges.

2.4 Virtual Grasping

Grasping is the primary and most frequent physical interaction technique people perform in everyday life (Feix et al., 2014b) and has been explored in VR/AR as a technical and computational challenge. Thus, researchers developed various grasp models using input devices that recognise the full hand such as wearables and cameras (as discussed in Section 2.2.2), focusing on developing algorithms for grasp recognition and intuitive grasp control in VR.

Early work of Ullmann and Sauer (Ullmann & Sauer, 2000) presented an approach to grasping virtual objects with a data glove. Their work aimed to develop a method that allows realistic grasping gestures which correspond to human grasping behaviour, to fulfil the demands of industrial applications in VR. Their algorithm used real-time collision detection and knowledge from real grasping behaviours where the grasping interaction depends on the behaviour of the object under the influence of gravity, the surface material of the object and the geometrical conditions at the point of contact between the object and the grasping hand. While this approach showed that users were able to grasp the virtual objects, the authors found that users had to be instructed to perform a specific grasp pose to trigger the interaction, which influenced the intuitiveness of the interaction. This led to the question *why was this grasping model developed based on real grasping knowledge?*, considering that whether or not users intuitively grasp virtual objects in the same way they grasp real objects has not been explored yet.

However, assumptions about this link between real and virtual grasping have been used as the basis knowledge in developing state-of-the-art virtual grasping interactions. Wan et al. (Wan, Luo, Gao, & Peng, 2004) developed grasp poses for virtual cubes, cylinders and spheres based on assumptions made from real grasping literature. A similar approach was followed by Valentini (Valentini, 2018) who developed three grasping poses that reflect main grasping poses used in real environments: cylindrical, spherical and pinch (see Figure 2.3). While this approach aims at enabling intuitive grasping in VR, by only allowing three grasp poses taken from real grasping patterns, the authors limit the interaction to predefined gestures that users need to learn to interact with the system, which negatively influences the naturalness of the interaction as discussed in Section 2.3.1. Moreover, the naturalness of the three grasp poses proposed based on real grasping literature can be questioned, considering that there is no evidence that users grasp virtual objects in the same way they grasp real objects. This is evident in the work of Jacobs and Froehlich (Jacobs & Froehlich, 2011) who developed a grasping model based on physics rules that mimic real grasping behaviour and found that the learning curve was rampant, as users realised they had to replicate grasping

(a) Grasp pose 1 (b) Grasp pose 2 (c) Grasp pose 3

Figure 2.3: Grasp poses developed by (Valentini, 2018) to reflect main grasping poses in reality.

movements from real environments instead of their intuitive virtual grasping approach (for example the majority of users initially tried to close their hands to form a fist instead of a realistic grasp).

While predefined grasping poses are common in virtual grasping, another important approach taken by HCI researchers is collision detection (Nasim & Kim, 2016). A collision detection algorithm calculates impact time by identifying two or more intersection points between the tracked hand and the virtual object (Borst & Indugula, 2005). Furmanek et al. (Furmanek et al., 2019) showed that while this approach can provide promising results for intuitive grasping, there is an increased perceptual uncertainty during target acquisition which led to slow interactions in VR. This uncertainty is often linked to the interaction overlap allowed by the detection algorithm, where an object might be too hard or too easy to grasp depending on the location where the collision is triggered. While there are limitations with the state-of-the-art algorithms involved in collision detection, the lack of haptic feedback might also be contributing to this uncertainty during grasping in VR. In real environments, contact is a very important event in human interaction with objects and surfaces, signalling task completion or ending two parallel processes of transport and grasp. However, when haptic feedback is missing in

Figure 2.4: Grasping accuracy analysis in immersive applications (Al-Kalbani et al., 2016a).

VR, users are guided in their grasps by visual perception only, which might introduce errors in estimating the location and bounds of the virtual objects, therefore influencing grasping behaviour (Zahariev & MacKenzie, 2007). Considering that providing haptic feedback is still a challenge in VR (Islam & Lim, 2022), and that an increased number of freehand interactions lack haptic feedback, this observation emphasizes the question discussed above: *why are state-of-the-art virtual grasping models mimicking real grasping behaviour, when VR environments are still challenged by providing realistic haptic feedback?*

Yet, researchers tried to mitigate this by using other types of feedback to replace natural haptic contact cues in grasping (Tinguy et al., 2019). For example, Zahariev and MacKenzie (Zahariev & MacKenzie, 2007) investigated auditory contact cues while grasping in VR by providing a sound when the hand was in contact with the virtual object. Their results showed that auditory feedback influenced the scaling of the movement, users adapting their grasp to the size of the cube, which was dependent of the availability of contact information. However, this led to slower interaction times due to participants slowly changing their grasp until they heard the auditory cue. This shows that when haptic feedback is replaced with other types of feedback, grasping behaviour changes, however, *how would this grasp behaviour change when no feedback to replace haptics is present?* The work presented in this thesis will address this question in more detail.

In an attempt to replicate real-world behaviour in VR, for improved presence and natural interactions, researchers explored ways to mitigate this lack of haptic feedback and provide a solution for hand penetration (where a hand sinks into virtual objects due to the lack of real physical constraints) (Borst & Indugula, 2005; Prachyabrued & Borst, 2014). Results showed that user experience was better in conditions that were not animating the hand to remain outside the object during interaction and instead provided users with accurate visual feedback of their hand movements. This shows that creating discrepancies between user interaction and visual feedback for mimicking real grasping in VR is not desirable. Alternatively, intuitive virtual grasping should be explored directly in VR to understand user behaviour when grasping virtual objects and what parameters influence their approach, to allow improvement of current systems that aim to provide intuitive virtual grasping interaction.

To address this, seminal work has started to explore virtual object grasping in a methodological way, analysing the accuracy and problems of freehand grasping in exocentric XR (Al-Kalbani et al., 2016a). Their work explored the viability of grasp categories with the preference of evaluating a VR medium grasp by defining the virtual representation of this grasp to be a visual holding of a virtual object without the use of haptic feedback and following a user-defined confirmation of the grasp. However, their work only explored interactions with cubes and spheres and only investigated interaction with one grasp. Additionally, the authors also analysed freehand grasping for dual visual feedback (Al-Kalbani, Williams, & Frutos-Pascual, 2016b), different grasp phases (Al-Kalbani, Frutos-Pascual, & Williams, 2017), virtual object visual effects such as shadows (Al-Kalbani, Frutos-Pascual, & Williams, 2019), however they did not explore other visual components or virtual object characteristics and their influence on grasping interaction patterns for a wider range of grasp types.

Considering this definition and framework of a virtual grasp, this thesis proposes the first VR Taxonomy of Grasp Types to explore different categorisations that can be given to virtual grasps for a wider range of virtual objects, visual cues and tasks, and structure these results to provide meaningful insights for grasping interaction design. The work presented in this thesis aims to address this by first understanding how grasping poses were studied and analysed in real environments which is presented in more detail in Chapter 3.

3 | Grasping Taxonomies

3.1 Introduction

Grasping is the primary and most frequent physical interaction technique people perform in everyday life (Holz, Ullrich, Wolter, & Kuhlen, 2008). It is defined as every static posture at which an object can be held securely with a single hand (Feix et al., 2009). As described in Chapter 2, virtual grasping has extensively been explored as a technical and computational challenge, however, when grasping virtual objects, users are often trained to use particular grasps with the design considerations and grasping constraints used in these solutions being applied from the body of knowledge available in real object grasping.

Grasping real objects has proved to be a demanding task (Supuk, Bajd, & Kurillo, 2011), determining a significant interest for understanding, studying and characterising aspects of human hand usage when interacting with objects (Redmond, Aina, Gorti, & Hannaford, 2010), especially in areas such as anthropology (Monaco, Sedda, Cavina-Pratesi, & Culham, 2014), hand surgery (Sollerman & Ejeskär, 1995), hand rehabilitation (Lukos et al., 2013) and robotics (Feix et al., 2009, 2014b; Bullock, Zheng, Rosa, Guertler, & Dollar, 2013; M. R. Cutkosky, 1989).

Chapter 2 showed that there is a need for understanding grasping patterns directly in VR, to understand how current intuitive VR systems can make use of intuitive and natural hand interaction. To analyse grasping patterns directly in VR, grasping metrics and existing methods for exploring grasp patterns need to be evaluated. Therefore, this chapter presents an overview of grasping work done in real environments, presenting an overview of grasping biomechanics, existing taxonomies and methods for using these taxonomies to explore interaction patterns for various influencing factors.

The chapter is structured as follows: Section 3.2 presents grasp definition, Section 3.3 presents grasping biomechanics, Section 3.4 presents grasp measures, Section 3.5 presents grasp taxonomies which details parameters of a grasp considered when analysing human grasping and Section 3.6 presents the most complete grasp taxonomy to date, which is currently further used in analysing grasp patterns for object characteristics and tasks.

3.2 Grasp Definition

The numerous skeletal and muscular degrees of freedom of the hand provide the human with an enormous dexterity that has not yet been achieved by any other species on earth (*Sensorimotor Control of Grasping: Physiology and Pathophysiology*, 2009). However, movement and function of the hand is not only a product of the internal degrees of freedom of the hand, but also the movement of the body and the arms as well as contact with the environment (Feix et al., 2016). The multitude of hand movements that can be performed by the hand can be divided into two main groups: *Prehensile*, or movements in which an object is seized and held partly or wholly within the compass of the hand; and *Non-prehensile*, or movements in which no grasping or seizing is involved but by which objects can be manipulated by pushing or lifting motions of the hand as a whole or of the digits individually (Napier, 1956).

Prehension, also known as the act of *grasping*, is the primary and most frequent physical interaction technique people perform in everyday life (Holz et al., 2008). A grasp is defined in the Oxford dictionary as:

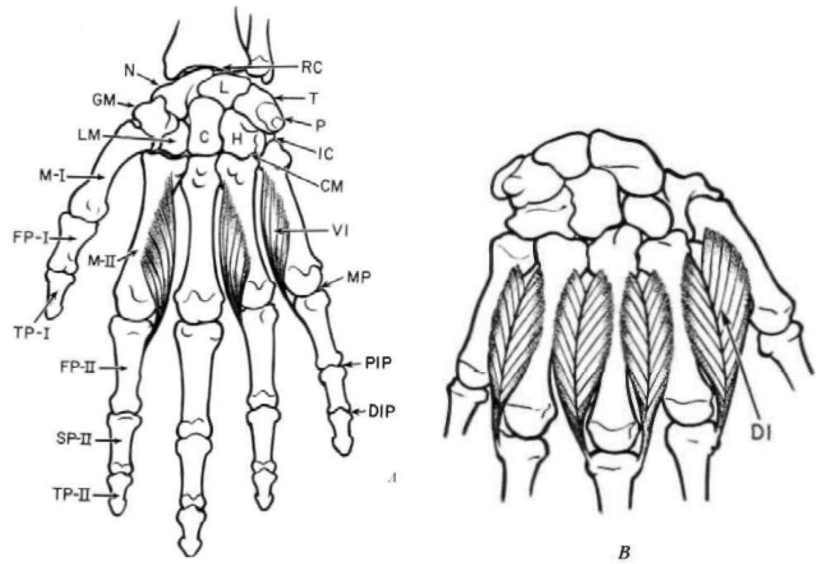
Definition 3.2.1 (Grasp) *A firm hold or a grip.*

However, researchers have defined grasping as the application of functionally effective forces by the hand to an object for a task, given numerous constraints (MacKenzie & Iberall, 1994), the act of relating finger positions and movements to a particular task (Lyons, 1985) or every static posture at which an object can be held securely with a single hand (Feix et al., 2009).

3.3 Grasping Biomechanics

The human hand represents a mechanism of the most intricate fashioning and one of great complexity and utility (Schwarz & Taylor, 1955). The hand consists of five digits, also known as fingers, that contain a collection of bones, tendons, muscles, ligaments, fascia and vascular structures. Overall, the human hand is composed of 27 bones and 39 muscles, with eight carpal bones in the wrist, five metacarpal bones in the palm, two phalanges in the thumb and three phalanges in each of the four fingers. The bones of the hand (Figure 3.1) naturally group themselves into the carpus, comprising eight bones which make up the wrist and root of the hand, and the digits, each composed of its metacarpal and phalangeal segments (Schwarz & Taylor, 1955).

Most of the muscles of hand and wrist lie in the forearm and, narrowing into tendons, traverse the wrist to reach insertions in the bone or ligamentous components of the hand. Figure 3.2 shows the main muscles of the hand and wrist: Abductor pollicis brevis (APB), Flexor digitorum sublimis (FDS), Flexor pollicis longus (FPL), Flexor digitorum profundus (FDP), Flexor pollicis brevis (FPB), Flexor carpi radialis (FCR), Palmaris longus (PL) and Flexor carpi ulnaris (FCU). The intrinsic muscles of the hand (those with both origin and insertion confined to wrist and hand) are, with the exception of the abductors of thumb and little finger, specialized for the adduction of the digits and for opposition patterns such as



(a) Volar view (b) Dorsal view

Figure 3.1: Bones and articulations of the hand showing Carpal Bones: Greater multangular (GM), Navicular (N), Lunate (L), Triquentrum (T), Pisiform (P), Lesser multangular (LM), Capitate(C), Hamate(H); Metacarpal bones: M-I, M-II, III, IV, V; First phalangeal series: FP-I, II, III, IV,V; Second phalangeal series SP-II, III, IV, V; Third phalangeal series TP-I, II, III, IV,V and Joints: Radiocarpal (RC), Intercarpal (IC), Carpometacarpal (CM), Metacarpophalangeal (MP), Proximal interphalangeal (PIP) and Distal interphalangeal (DIP) (Schwarz & Taylor, 1955).

making a fist or a spherical grasp (Schwarz & Taylor, 1955).

These are all used together to mediate dexterous postures and interactions that are being performed everyday by humans, such as grasping (MacKenzie & Iberall, 1994) which has proved to be a demanding and complex task (Supuk et al., 2011). Moreover, while the human hand is a very complex tool, the size of the human hand and its bones and muscles is relatively small (MacKenzie & Iberall, 1994), allowing a higher bandwidth for mobility, adaptability and control that allows the hand to perform both small and large deformations when required (Kristan et al., 2000). Moreover, due to its complex mechanical design, the human hand proposes

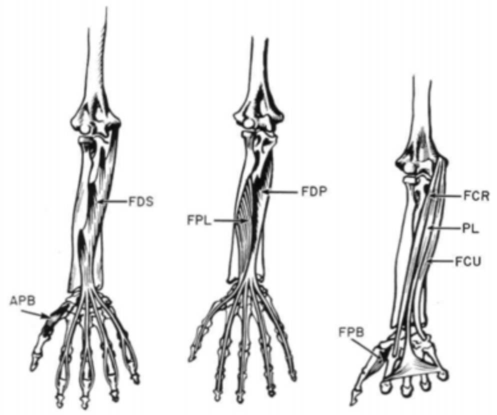


Figure 3.2: Flexors of wrist and digits: Abductor pollicis brevis (APB), Flexor digitorum sublimis (FDS), Flexor pollicis longus (FPL), Flexor digitorum profundus (FDP), Flexor pollicis brevis (FPB), Flexor carpi radialis (FCR), Palmaris longus (PL) and Flexor carpi ulnaris (FCU) (Schwarz & Taylor, 1955).

28 degrees of freedom, allowing high dexterity for grasping objects of different sizes and shapes, adjusting depending on the intended task (K. M. B. Bennett & Castiello, 1994).

3.4 Grasp Measures

Derived from the complexity and physiology of the human hand, researchers focused on observing and classifying grasping movements (M. R. Cutkosky, 1989), aiming to introduce simplifications of the analysis process to allow a better understanding of the human grasping capabilities for anthropology, hand surgery, hand rehabilitation, robotics, developmental psychology and virtual environments (MacKenzie & Iberall, 1994).

To achieve this, researchers defined several analytical grasp quality measures for describing a successful grasp. Cutkosky and Kao (M. Cutkosky & Kao, 1989) proposed compliance as an important measure for describing a grasp, focusing on the effective compliance (inverse of stiffness) of the grasped object with respect to the hand. It refers to a function of grasp configuration and structural compliances in the links, joints and fingertips (M. R. Cutkosky, 1989). Mason et al. (Mason & Salisbury, 1985) proposed connectivity, which refers to the number of independent parameters needed to completely specify the position and orientation of the object with respect to the palm. Another metric for measuring grasp performance is force closure, which, assuming that external forces maintain contact between the fingers and the object, looks at the ability of the object to move without slipping when the finger joints are locked (Ohwovoriole & Roth, 1981). From this, the idea of assessing form closure emerged, which looks at the ability of the grasp to hold an object when external forces are applied from any direction (Mason & Salisbury, 1985). Kerr et al. (Kerr & Roth, 1986) looked at grasp isotropy, which measures if the grasp configuration allows the finger joints to accurately apply forces and movements to the object. For example, if one of the fingers is nearly in a singular configuration, it will be impossible to accurately control force and motion in a particular direction.

Other measures have been proposed to assess the quality of a grasp such as manipulability (Kerr & Roth, 1986) and resistance to slipping (M. R. Cutkosky & Wright, 1986; Kerr & Roth, 1986), however the metric that encompasses the main grasp measures that ensure a successful grasp is grasp stability (Pollard & Lozano-Perez, 1990) and dexterity (MacKenzie & Iberall, 1994). Stability refers to the ability of the grasp to return to its initial configuration after being disturbed by an external force (M. R. Cutkosky, 1989; MacKenzie & Iberall, 1994) while dexterity refers to how accurately the fingers can impart larger motions or forces, and sensitivity or how accurately fingers can sense small changes in force and position (MacKenzie & Iberall, 1994).

While human hands perform stable grasps with ease during everyday tasks, due to the mechanical complexity of the hand, understanding the patterns of stable grasps and the parameters that influence these grasping approaches is a challenging task (Supuk et al., 2011). Therefore, researchers started to focus on understanding how a stable grasp is achieved naturally by the human hand and classified hand postures based on task or finger positions in regards to the object which were then organised in grasp taxonomies (MacKenzie & Iberall, 1994).

3.5 Grasp Taxonomies

Taxonomies, defined as the "science of classification" (Bowman & Hodges, 1999) have been highly used for classifying grasping patterns to provide a deep understanding of the way humans grasp objects, being an important contribution in many domains ranging from anthropology, medical literature, rehabilitation, psychology and robotic arm design among many others (Feix et al., 2016). While it is clear that the complexity and variety of uses of the human hand makes the categorisation and classification of grasps a challenging task, researchers attempted to simplify this process by identifying grouping mechanisms based on independent parameters that might influence grasping approach. For example, a now classic taxonomy proposed by Schlesinger (Schlesinger, 1919) defined hand postures based on the shape of the object to be grasped. His work captured the versatility of human hands for designing functionally-effective prosthetic hands and focused on determining what specific functionality was needed for grasping and holding various objects (e.g. book, pen, matchbox) and devised a minimum set of six grasp postures: Cylindrical for cylindrical objects, Tip for very small objects, Hook for heavy objects, Palmar for flat thick objects, Lateral for flat thin objects and Spherical for spherical objects (see Figure 3.3). Keller et al. (Keller & Zahm, 1947) used Schlesinger's taxonomy (Schlesinger, 1919) for identifying a logical basis for defining these patterns and stated that the object-contact pattern furnishes a satisfactory basis for grasp classification. Therefore, based on photographic observation of humans picking up and holding common objects used in everyday life, they selected three common grasp types from the ones originally proposed by Schlesinger (Schlesinger, 1919): palmar, tip and lateral.

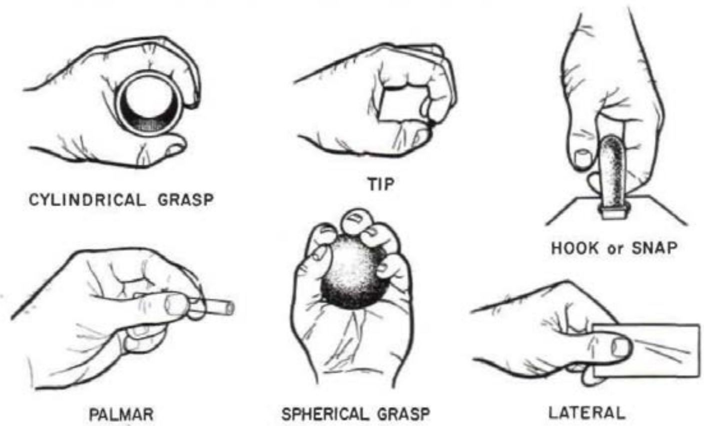


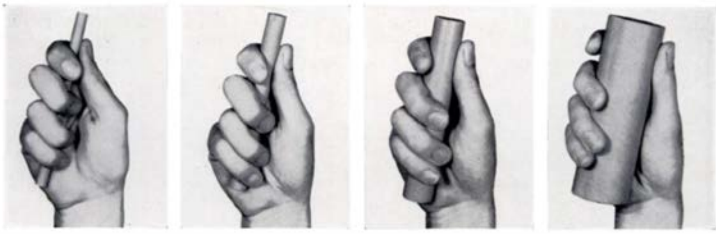
Figure 3.3: Grasp posture classification proposed by Schlesinger (Schlesinger, 1919): Cylindrical grip for cylindrical objects, Tip for very small objects, Hook for heavy objects, Palmar for flat thick objects, Lateral for flat thin objects and Spherical for spherical objects as presented in (MacKenzie & Iberall, 1994).

A similar approach was followed by Griffiths (Griffiths, 1943) who divided the functions of the hand into five main grips: Cylinder , Ball, Ring, Pincer and Pliers which was later simplified by Slocum and Pratt (Slocum & Pratt, 1946) to only show 3 main grasps, by focusing on the hand opposition parts in relation to the fingers: Grasp, a coupled action between the fingers and the opposite palm and thumb of the hand, Pinch, thumb pad against pads of the opposing fingers and Hook, flexed fingers where their pads are parallel and marginally away from the palm. Further, McBride (McBride, 1942) stressed the importance of the hand surfaces being considered for hand posture classification and proposed grasping by the hand as a whole, grasping between the thumb and the fingers and grasping by the combined use of the palm and the digits (as prehension patterns) as the most common grasping patterns.

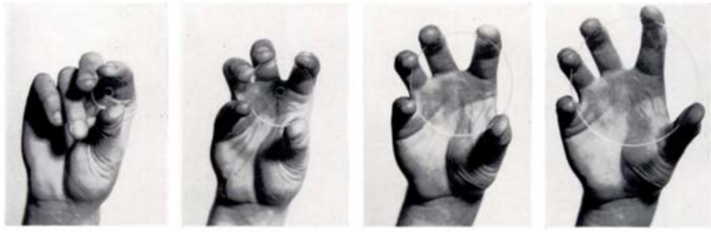
3.5.1 Power and Precision Classification

While the taxonomies presented above provide an insightful and extensive overview of grasp postures, Napier (Napier, 1956) argued that they are not clearly defined nor are they comprehensive, as they represent a series of functional end-results rather than a fundamental analysis of the potential hand as a whole (Griffiths, 1943), they do not consider the numerous functions of the hand in which the thumb is not in opposition (Slocum & Pratt, 1946) or are somewhat arbitrarily conceived and do not have any particular functional or anatomical basis (McBride, 1942).

To address this, he stressed the importance of developing a fundamental approach to the problem of the function of the hand as a whole to provide an effective and generally acceptable terminology, and proposed two discrete patterns of movement from both the anatomical and functional points of view: precision grip and power grip. He defined power grips as the grips where the object is held in a clamp formed by the partly flexed fingers and the palm, counter pressure being applied by the thumb lying more or less in the plane of the palm and precision grips as the grips where the object is pinched between the flexor aspects of the fingers and the opposing thumbs (see Figure 3.4). He provides a detailed overview of power and precision grips showing that grasp stability (see Section 3.4) is a fundamental requisite of all types of prehension, meaning that the object, whether it is fixed or freely movable, should be held securely. Further, he states that the power and precision patterns either separately or in combination, provide the anatomical basis for all prehensile activities, whether skilled or unskilled and claims that these two embody the whole range of prehensile activity of the human hand.



(a) Power grips



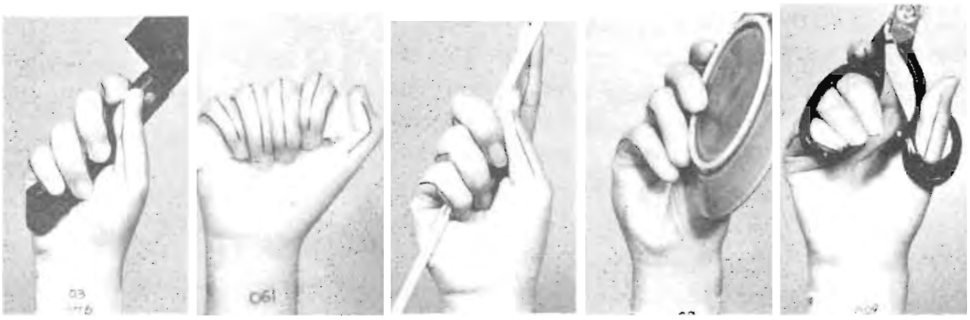
(b) Precision grips

Figure 3.4: Examples of power and precision grips from Napier's work (Napier, 1956)

However, Landsmeer (Landsmeer, 1962) argued that using power and precision to define grasps in a dynamic as well as static sense might not be valid, noting that the dynamics of gripping produce a particular grip, and the static concept indicates the final state of gripping, showing the need for further analysis on whether both forms of Napier's grips may be regarded as a variant of the final state, or whether the movements leading to this final state may be taken in both cases as variants of an equivalent movement pattern. To address this, he provides a detailed overview of the two prehension concepts in the context of both dynamic and static stages of prehension and argues that: In the case of power grip, in the dynamic phase, the hand must be open, and the fingers and thumb must take up a position suitable for grasping the object and further transition to the static phase (the actual grip). In the precision grip there is no question of a static phase, since the fingers themselves manipulate the object, and there is no point in distinguishing dynamic and static phases in this movement pattern. He therefore suggests that the term "grip" is not applicable in this scenario and therefore suggested that this manoeuvre to be called "handling", introducing the term "precision handling", which is described as the posture where the fingers are first placed to hold the object, being followed by the motion of fingers in relation to each other, which perform the actual handling or manipulation. This revision highlighted the importance of considering different stages of grasping action and led to researchers focusing on understanding human grasping in both dynamic and static phase. Based on this, Patkin (Patkin, 1981) introduced dynamic grasps that represent the hand when it can still act while grasping an object, such as writing with a pen or cutting with scissors: external precision grasp, internal precision grasp and double grip.

Kamakura et al. (Kamakura et al., 1980) also considered Landsmeer's analysis (Landsmeer, 1962) and focused on investigating grasping in the static phases only. They also reviewed Napier's classification (Napier, 1956) and argued that for clinical settings, a more detailed classification is necessary when positions are stressed in evaluation and re-education of disabled hands. They agree with Napier's classification and present grasp categories that are based on the nature of the task (Napier, 1956) and argue that the position of the fingers can also determine the characteristics of grasping patterns and therefore introduce grasp types as part of these grasp categories. They present power grip category which is defined as the grasp where a wide area of hand, including part of the palm, makes contact with the object and includes five grasp types: standard type (PoS, Figure 3.5 a) hook type (PoH, Figure 3.5 b), index finger extension type (PoI, Figure 3.5 c), grip extension type (PoE, Figure 3.5 d) and distal type (PoD Figure 3.5 e).

While Landsmeer's work suggested that precision grasping (handling) refers to dynamic grasps, Kamakura et al. (Kamakura et al., 1980) present precision category as part of their analysis of static grasps. The category includes five grasp types: parallel mild flexion grip (PMF Figure 3.6 a), surrounding mild flexion grip



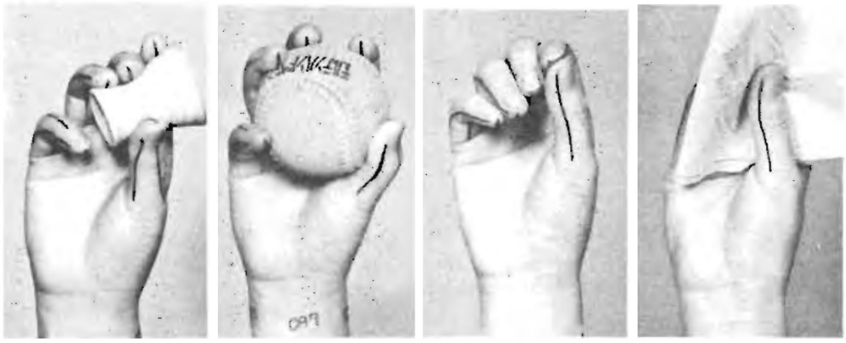
(a) PoS (b) PoH (c) PoI (d) PoE (e) PoD

Figure 3.5: Power grasp types classified in three grasp categories: Power, Precision and Intermediate as described by Kamakura et al. (Kamakura et al., 1980): Pos - Standard Type, PoH - Hook Type, PoI - Index Finger Extension, PoE - Extension Type, PoD - Dystal Type.

(SMF Figure 3.6 b), tip prehension (Tip Figure 3.6 c) parallel extension grip (PE Figure 3.6 d).

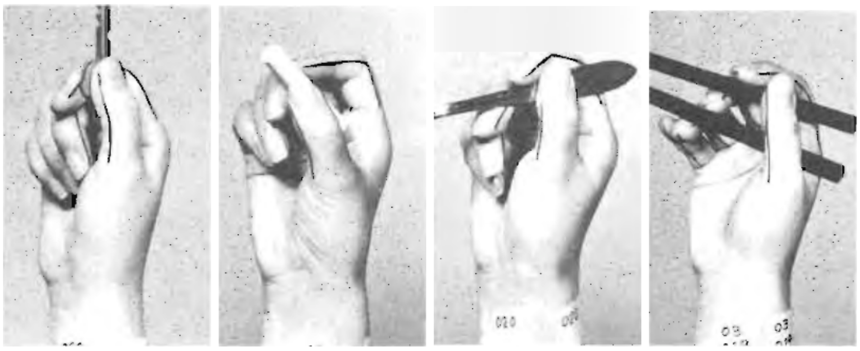
Napier mentioned the possibility of a combined grip in his classification (Napier, 1956), however his definition of a combined grip referred to grips where the hand is required to use a power and precision grip at the same time (two different grasps are performed at the same time with the same hand), the grasps being still categorised as power or precision. Kamakura et al. (Kamakura et al., 1980) introduced Intermediate grasps, where the hand is in intermediate position between power grip and precision grip and include four grasp types: lateral grip (Lat Figure 3.7 a), tripod grip (Tpd Figure 3.7 b), tripod variation 1 (TV1 Figure 3.7 c) and tripod variation 2 (TV2 Figure 3.7 d).

A similar approach was followed in (Bendz, 1974) where the authors focused on supplementing Napier's classification and suggested Flexion Grip and Extension Grip as subtypes of both the power and the precision grips, however they did not mention how to differentiate between these two grasp types. While previous taxonomies presented common posture names (Schlesinger, 1919), described pat-



(a) PMF (b) SMF (c) Tip (d) PE

Figure 3.6: Precision grasp types classified in three grasp categories: Power, Precision and Intermediate as described by Kamakura et al. (Kamakura et al., 1980): PMF - Parallel Mild Flexion Grip, SMF - Surrounding Mild Flexion Grip, Tip - Tip Prehension, PE - Parallel Extension



(a) Lat (b) Tpd (c) TV1 (d) TV2

Figure 3.7: Intermediate grasp types classified in three grasp categories: Power, Precision and Intermediate as described by Kamakura et al. (Kamakura et al., 1980): Lat - Lateral Grip, Tpd - Tripod Grip, TV1 - Tripod Variation 1, TV2 - Tripod variation 2.

terns based on task (Napier, 1956) or analysed the terminology for various grasping phases (Landsmeer, 1962), Kamakura et al. (Kamakura et al., 1980) used the grasp types proposed to further link them to specific objects, to provide a clear and structured taxonomy of grasp types (See Figure 3.8).

Category	Pattern	Objects
Power Grip	PoS	Knife, kitchen knife, handle of a pan, can opener (stick type), wrench (210 g), handle of a file, handle of a magnifying glass (10 mmϕ, 90 mmL), pair of Japanese scissors, umbrella
	PoH	Bucket (with 4 kg load inside)
	PoI	Fork,* knitting needle, tracing spatula (for moderate trace)
	PoE	Saucer (with a cup on top, 148 mmϕ), can opener (flat type)
	PoD	Pair of scissors
Intermediate Grip	Lat	(none)
	Tpd	Chalk
	TV1	(none)
	TV2	Pair of chop sticks
Precision Grip	PMF	Sake cup (43 mmϕ, 33mmϕ, 48mmH), match stick, abacus on a table,* opera glasses
	SMF	Ball (70 mmϕ)
	Tip	Zipper tag of pants, drawing pin, needle (for threading)*
	PE	Toilet paper
Grip Involving No Thumb	Add	Cigarette (smoking), cigarette (holding)

Figure 3.8: Taxonomy of grasp types connected to objects as presented in (Kamakura et al., 1980)

A different approach to creating subtypes of the power and precision grasp categories defined by Napier (Napier, 1956) and defining intermediate categories was followed by Lyons (Lyons, 1985). He defined a set of three grasps: Encompass Grasp, which represents a power grasp and shows the hand completely enveloping the object being grasped; Lateral grasp, which is defined as a pinch grasp which is intended to be used for precise manipulation of long objects, as opposed to Kamakura et al. (Kamakura et al., 1980)'s definition of a lateral pinch describing an intermediate grasp representing small, flat objects being grasped between the lateral aspect of the middle or distal phalanx of the index finger and the pulp of the thumb (see Figure 3.7); and Precision grasp which is defined as a general grasp where the grasped object is held between the finger tips and allows the maximum amount of possible manipulation. However, unlike Kamakura et al. (Kamakura et al., 1980) who proposed subtypes of grasp categories based on finger positions during grasping, Lyons particularized each of these categories to suit the size and shape of the target object. For example, the distance between the thumb and the fingers in a lateral grasp will be a function of the anticipated object size, suggesting that these three grasps will remain the same with the measurement to be changed being grasp aperture only. Similar to Kamakura et al. (Kamakura et al., 1980) he also emphasizes the need for defining intermediary grasps between power and precision, however, instead of proposing an Intermediate category with subtypes, Lyons defined a metric Grasp Index (GI) that shows the firmness and precision required for a grasp and presents four possible categories: No Precision, No Firmness (NP, NF); Precision, No Firmness (P, NF); No Precision, Firmness (NP, F) and Precision, Firmness (P,F) and presents these categories linked to objects of different shapes and sizes as shown in Figure 3.9.

In the robotics literature, Cutkosky and colleagues (M. Cutkosky & Wright, 1986; M. R. Cutkosky & Howe, 1990; M. R. Cutkosky, 1989) also extended Napier's precision and power classification by further classifying power grasping into nine grasp subtypes and precision grasping into seven subtypes using a set of grasp attributes. In power grasps, the emphasis is on grasp stability and security while in precision grasps the emphasis is on dexterity (see Section 3.4). The complete grasp taxonomy is shown in Figure 3.10.

Functionality	Size	Shape	GRASP
(NP,NF)			
	Small	Flat	ENC.
		Round	ENC.
	Long	Flat	LAT.
		Round	LAT
	Large	Flat	LAT.
		Round	ENC.
(P,NF)			
	Small	Flat	PRE.
		Round	PRE.
	Long	Flat	PRE.
		Round	PRE.
	Large	Flat	LAT.
		Round	ENC.
(NP,F)			
	Small	Flat	ENC.
		Round	ENC.
	Long	Flat	LAT.
		Round	ENC.
	Large	Flat	LAT.
		Round	ENC.
(P,F)			
	Small	Flat	PRE.
		Round	PRE.
	Long	Flat	LAT.
		Round	LAT.
	Large	Flat	LAT.
		Round	ENC.

Figure 3.9: Taxonomy of grasp types categorised based on Grasp Index (GI) showing four categories: No Precision, No Firmness (NP, NF); Precision, No Firmness (P, NF); No Precision, Firmness (NP, F) and Precision, Firmness (P,F) and the three grasps presented: Encompass Grasp (ENC); Lateral Grasp (LAT) and Precision Grasp (PRE) as described by (Lyons, 1985)

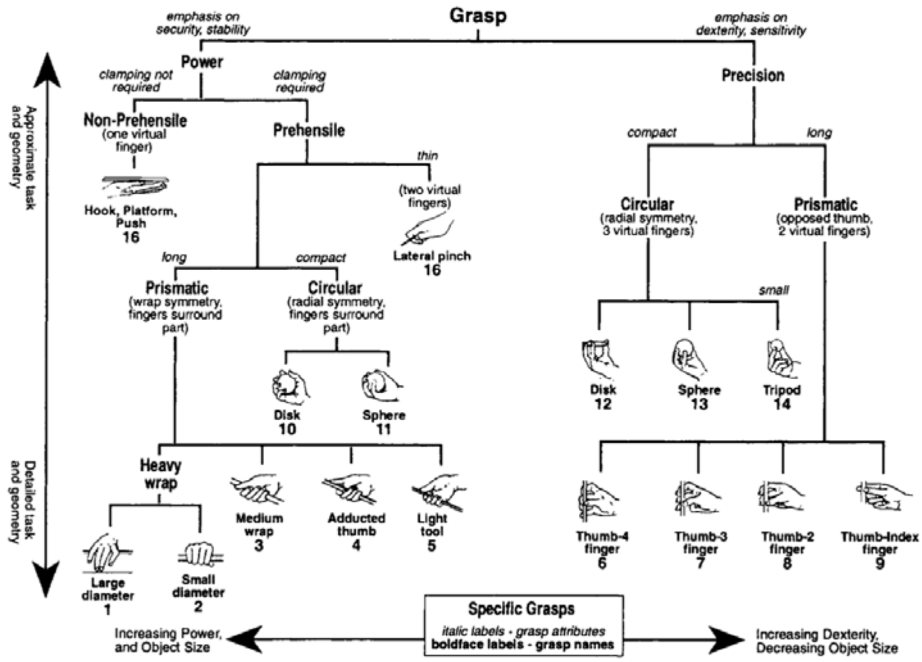
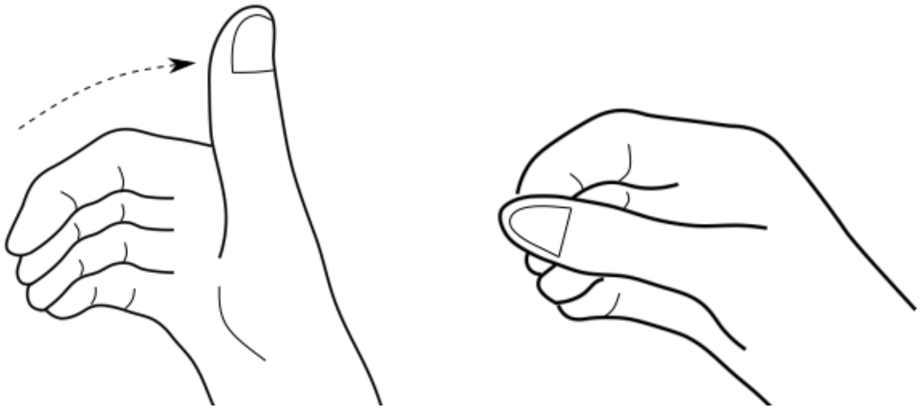


Figure 3.10: Taxonomy of grasp types categorised by Power and Precision grasp categories with subtypes

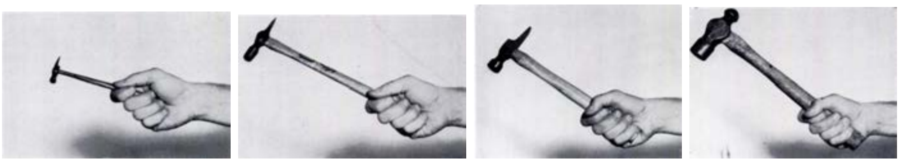
3.5.2 Thumb Adducted/Abducted Classification

Napier (Napier, 1956) noted that the level of precision in a power grasp is dependent on the position of the thumb and therefore introduces the analysis of the posture of the thumb in two patterns: abducted and adducted (see Figure 3.11), where the thumb fulfils the need for stability. He showed that in the power grip the thumb is adducted at both metacarpo-phalangeal and carpo-metacarpal joints while in the precision grip the thumb is abducted at both these joints. In general, he claims that the greater the force required of the grip as a whole, the more the thumb is required to act as a reinforcing mechanism and less it is able to contribute to precision as seen in Figure 3.12. Kamakura et al. (Kamakura et al., 1980) also introduced the Adduction category where the thumb is not involved and defines a grasp where a small, light object is held between adjacent fingers. Other uses



(a) Thumb Adducted (b) Thumb Abducted

Figure 3.11: Thumb positioning in grasping recognition according to the Human GRASP Taxonomy for grasping real objects (Feix et al., 2016)



(a) A (b) B (c) C (d) D

Figure 3.12: A series of hammer grips demonstrating the changing relationship of the thumb to the shaft of the hammer as the size of the tool increases. Subfigure A presents a pin hammer, Subfigure B presents a Warrington hammer, Subfigure C presents a cross-pein hammer and Subfigure D presents a ball-pein hammer (Napier, 1956)

of this posture are seen when small objects need to be removed from tight places, such as grasping coins from pockets (MacKenzie & Iberall, 1994), which Napier defined as scissors grip while Kapandji (Kapandji, Honoré, & Poilleux, 1982) anatomically described it as the interdigital latero-lateral grip.

3.5.3 Opposition and Virtual Finger Classification

Another way to analyse what the hand is doing in prehension is to focus on the fact that a posture involves at least two forces being applied in opposition to each other against the object's surfaces (MacKenzie & Iberall, 1994). Iberall et al. (Iberall, Bingham, & Arbib, 1986) used the term opposition to describe three basic directions along which the human hand can apply forces. Based on the theory that prehensile postures are constrained by the way the hand can apply opposing forces around an object for a given task, they propose a grasp classification method based on three types of opposition: pad opposition, palm opposition and side opposition as shown in Figure 3.13.

Pad opposition occurs between hand surfaces along a direction generally parallel to the palm. This usually occurs between volar surfaces of the fingers and thumb near or on the pads. Palm opposition occurs between hand surfaces along a direction generally perpendicular to the palm and Side opposition occurs between hand surfaces along a direction generally transverse to the palm (MacKenzie & Iberall, 1994). Using this theory, Iberall et al. (Iberall et al., 1986) proposed that existing classifications can be reanalysed to consider the types of opposition in defining grasp types. For example, they show that the basic precision grasp proposed by Iberall and Lyons (T. Iberall, 1984) which involved the separate use of the index finger from the middle finger can now be explained as pad opposition between the thumb and the index finger, in conjunction with side opposition between the thumb and the middle finger. Their grasp taxonomy provides mapping of existing taxonomies such as the work of Schlesinger (Schlesinger, 1919), Napier (Napier, 1956), Iberall and Lyons (T. Iberall, 1984) and Cutkosky and Wright (M. Cutkosky & Wright, 1986) to show types of opposition for every posture proposed in previous taxonomies.

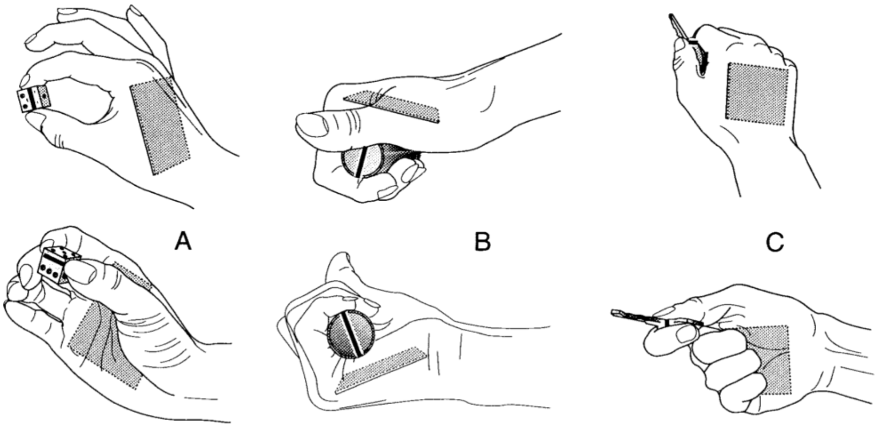


Figure 3.13: Grasping postures consist of combinations of three basic ways the hand can provide oppositions around objects. The solid line shows the opposition vector seen in the object. The shaded area represents the plane of the palm. A. Pad opposition which occurs along an axis generally parallel to the palm; B. Palm opposition which occurs along an axis generally perpendicular to the palm and C. Side opposition which occurs along an axis generally transverse to the palm (MacKenzie & Iberall, 1994)

.

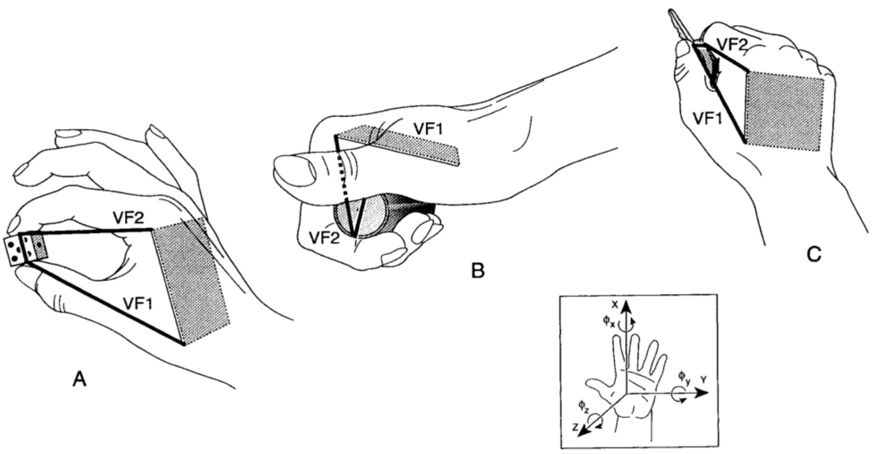


Figure 3.14: Oppositions can be described in terms of virtual fingers, relative to a hand coordinate frame placed on the palm. A shows pad opposition, B shows palm opposition and C shows side opposition as described by (Iberall, 1987)

In an observation study focusing on how users grasp different sized mugs, Arbib et al. (Arbib, 1985) noted that the length of the mug handle influenced the number of fingers used for the grasp. However, they show that the task remained the same: a finger was placed on top of the handle, one or more fingers were placed inside the handle and if available, fingers were placed against the outside of the body (MacKenzie & Iberall, 1994). Based on this, they suggested that each of these functions were being performed by a virtual finger (VF) as the method for applying the force. A virtual finger is an abstract representation, a functional unit for a collection of individual fingers and hand surfaces applying an oppositional force. Real fingers group together into a virtual finger to apply force or torque opposing other VFs or task torques (MacKenzie & Iberall, 1994).

Figure 3.14 shows that two virtual fingers apply forces in opposition to each other and the direction of these forces is relevant in order to clarify the type or types of oppositions: pad opposition occurs (Figure 3.14 A) along the X axis between the thumb as VF1 and one or more fingers as VF2. Palm Opposition (Figure 3.14 B) occurs along the Z axis between the palm as VF1 and the fingers as VF2. Side opposition (Figure 3.14 C) occurs along the Y axis between the thumb as VF1 and the index finger as VF2.

Opposition and virtual fingers can be identified in existing grasp types presented in the above-mentioned taxonomies; Palm opposition is used in the coal hammer grasp (Napier, 1956), cylindrical grasp and spherical grasp (Schlesinger, 1919) where the palm (VF1) is in opposition to the fingers and thumb (VF2). Pad opposition is used in precision grasps such as palmar pinch and tip prehension (Kamakura et al., 1980), where the thumb (VF1) is in opposition to one or more fingers (VF2). Side opposition is used in lateral prehension, between the thumb (VF1) and the radial side of the fingers (VF2) as well as in the adduction grip (Kamakura et al., 1980) where one digit (VF1) is in opposition to another digit

Authors	Grasp Terminology
(Schlesinger, 1919)	cylindrical, tip, hook, palmar, spherical, lateral
(McBride, 1942)	whole hand grasp, palm, digits grasp, thumb, finger grasp
(Griffiths, 1943)	cylinder grip, ball grip
(Slocum & Pratt, 1946)	grasp, pinch, hook
(Napier, 1956)	power grip, precision grip, combined grip, hook grip
(Landsmeer, 1962)	power grasp, precision handling
(Kamakura et al., 1980)	power-grip standard, power grip-index ext, power grip-extension, parallel mild flexion grip, tip prehension, surrounding mild flexion grip, parallel extension grip, tripod grip, tripod grip-var 1, tripod grip-var 2, lateral grip, power grip-hook, power grip-distal, adduction grip
(M. R. Cutkosky & Howe, 1990)	large diameter, small diameter, medium wrap, adducted thumb, light tool, thumb-4 finger, thumb-3 finger, thumb-2 finger, thumb-index finger, disk, sphere, tripod, lateral pinch, hook, platform push

Table 3.1: Overview of grasp terminology proposed by researchers in grasp taxonomies.

(VF2) (MacKenzie & Iberall, 1994),

3.6 The GRASP Taxonomy of Human Grasp Types

3.6.1 Rationale

The majority of grasp taxonomies have been developed with the goal of understanding what types of grasps humans commonly use in everyday tasks and were further used as an inspiration for designing robotic and prosthetic hands. However, due to the complexity and variety of uses of the human hand, the classification and categorisation of the hand function is a challenging task (Feix et al., 2016) which led to a lack of consensus in defining the terminology of a range of grasp types that humans commonly use. Table 3.1 shows an overview of the proposed grasp terminology by researchers, showing the variety of grasp names proposed in the literature. To allow a better understanding of the human hand and create a framework for investigating human hand use for various tasks and objects, Feix et al. (Feix et al., 2009) emphasized the need for creating a common terminology of grasp types to allow further investigation into human hand use.

3.6.2 Terminology

To create this common terminology and develop a complete grasp taxonomy, Feix et al. (Feix et al., 2009) reviewed 22 existing grasp taxonomies from literature ranging from the field of robotics, developmental medicine, occupational therapy and biomechanics. They reviewed grasping classes defined by researchers and provided an overview for categorising static grasps, also known as grasps where the object is in constant relation to the hand (Landsmeer, 1962). They reviewed the methodology of categorising grasps in power, precision and intermediate grasps (see Section 3.5.1), the basic directions relative to the hand coordinate frame in which the hand can apply forces on the object to hold it securely, also known as opposition types (see Section 3.5.3), thumb adduction and abduction (see Section 3.5.2) and the way several fingers work together as a functional unit, known as virtual finger (see Section 3.5.3). They grouped similar grasps together and presented them in a matrix, where columns are arranged according to Power, Intermediate and Precision requirements, with the next finer differentiation being the type of opposition (Palm, Pad or Side Opposition). The opposition type is also defining the VF1: in the case of Palm Opposition, the palm is mapped into VF1; in Pad and Side Opposition, the thumb is VF1. Then, the position of the thumb is used to differentiate between the two rows. Figure 3.15 shows this taxonomy that describes 33 grasp types, more than the previously developed taxonomies of Kamakura et al. (Kamakura et al., 1980) (14 grasps), Cutkosky (M. R. Cutkosky & Howe, 1990) (16 grasps) or Kapandji et al. (Kapandji et al., 1982) (21 grasps), being the most complete grasp taxonomy to date.

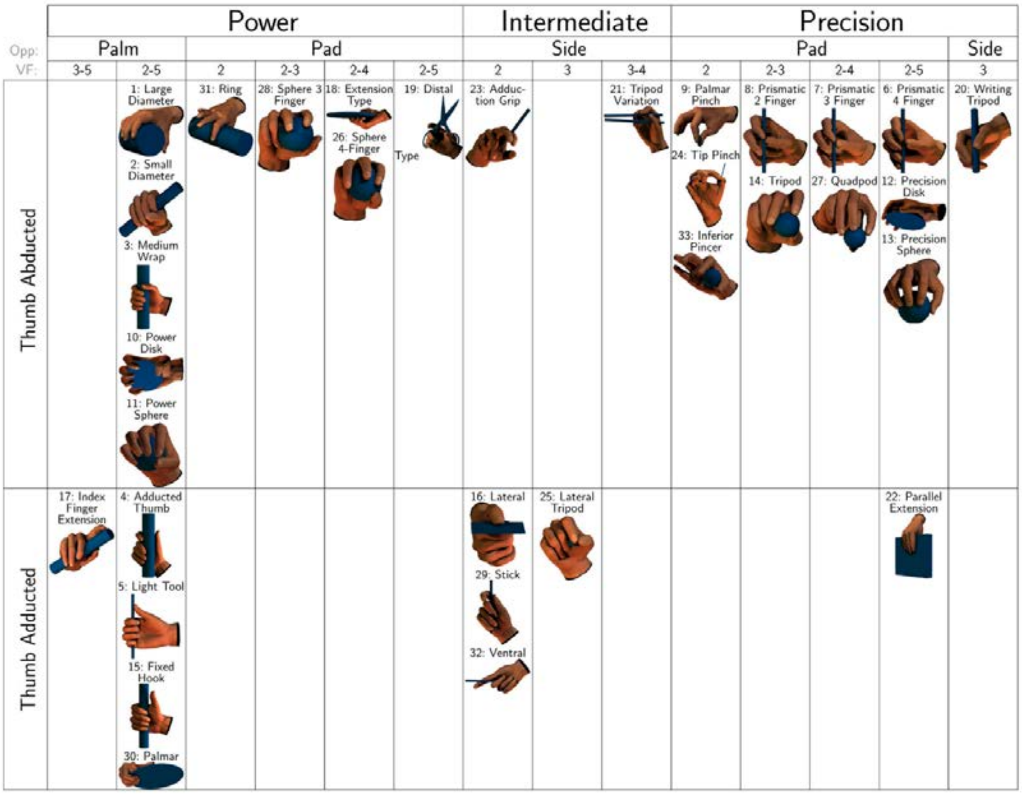


Figure 3.15: The most complete Human GRASP Taxonomy to date, presented in (Feix et al., 2009).

3.6.3 Taxonomy Use

Being the most up to date and complete grasp taxonomy, researchers further used it to analyse grasping patterns in the presence of different parameters, to create more complete grasp taxonomies based on influencing parameters. Bullock et al. (Bullock et al., 2013) used the GRASP taxonomy as a basis in finding small sets of versatile human grasps. For this, they analysed 19 hours of video with over 9000 grasp instances from two housekeepers and two machinists and used the terminology provided in the GRASP taxonomy to tag the grasps collected. Further, they analysed the number of times each grasp was used for each profession and provided an overview of the most common grasps used by machinists and housekeepers.

Feix et al. (Feix et al., 2014b) also emphasized the need for investigating the properties of objects humans interact with on a daily basis and correlating them to grasp choices (from the GRASP taxonomy) for a better understanding of grasping patterns in real environments. This assumption that object characteristics influence grasping patterns has served as basis for the development of the first grasp taxonomies (Schlesinger, 1919; Kamakura et al., 1980), however, a detailed analysis of object characteristics has not been conducted before. To address this, Feix et al. (Feix et al., 2014b) conducted a user study where they focused on categorising grasps and objects and correlating them for a better understanding of their relationship. For this, they recorded two machinists and two housekeepers while wearing a head-mounted camera during eight hours of professional work. Two raters then tagged the right-handed grasps in the video, using the common GRASP terminology (Feix et al., 2009). They showed that object characteristics and task influence grasping patterns in real environments, providing recommendations for improving robotic grasping models.

They connected the defined object properties, to grasp types and categories (from the GRASP taxonomy of human grasp types (Feix et al., 2009)) and developed a grasp taxonomy, showing that the most common grasp type for all objects is medium wrap, which is consistent with the finding of (Bullock et al., 2013). However, when looking at individual object categories, grasping patterns vary for both grasp types and grasp dimension. Cylinder-shaped objects are generally grasped with a medium wrap while small and lightweight objects are generally grasped with the lateral pinch. Moreover, long objects are usually grasped from the side, rather than using the longest dimension. A disk is usually grasped from the top, whereas a short prism is grasped by the smallest dimension and irregular objects unsurprisingly show the largest variation in the grasped dimension. These findings indicate that object properties influence grasping patterns in real environments and therefore emphasize the importance of correlating grasp patterns to object characteristics for a deep understanding of human grasping approach. Moreover, they show that their proposed methodology of analysing grasp category and grasp type against influencing parameters such as object characteristics provides detailed and insightful information on grasping patterns in real environments.

Feix et al. (Feix, Bullock, & Dollar, 2014a) extended this taxonomy by focusing on how strongly different properties of the task influence grasp choice and how well those properties can predict the grasp chosen. For this, they assigned each task a set of properties that they hypothesized were important for grasping and that could also be easily assigned by visual observation: constraints, which describe the degrees of freedom and nature of the constraints of the grasped object; functional class, which describes at a high-level what is being done with the object; and force, which describes whether the force being applied is to lift the object or based on some additional task property such as opening a door. Using the same recording of the two machinists and housekeepers (Feix et al., 2014b), the authors tagged the tasks and the corresponding grasp types to extend the grasp taxonomy. Their results show that task influences the choice of grasp, presenting patterns of grasp choice for independent tasks (for example the Ring grasp is primarily used for "turning knob" and "turning handle"). Their approach has shown to provide useful information about the objects that people interact with, the tasks people perform and the grasps they use. This information is useful for defining performance specifications and testing conditions for robotic hands, providing basic heuristics for developing grasp planners, targeting rehabilitation efforts toward essential hand functionality and aiding in the design of haptic interfaces or other devices that should interact with the hand in a natural manner (Feix et al., 2014a).

Therefore, the work of Feix et al. including the definition of a common terminology of grasp types (Feix et al., 2009) as well as the analysis of object characteristics (Feix et al., 2014b) and task properties (Feix et al., 2014a) in correlation with grasp patterns is an important contribution to the grasping community, especially for developing robotic grasp models (Elangovan, Chang, Gao, & Liarokapis, 2022; Liu, Jiang, Liu, & Ming, 2022), being the state of the art method in developing grasp taxonomies for a deep understanding of grasping patterns in real environments. Therefore, the work presented in this thesis will use this method as a basis for developing the first VR grasping taxonomy, adapting it based on VR requirements and limitations, which are described in more detail in Chapter 4.

4 | A Method for the First VR Taxonomy of Grasps

4.1 Introduction

Analysing grasping patterns directly in VR is key to revealing user behaviour and informing design decisions when developing intuitive grasp interactions, as emphasized in Chapter 2. While virtual grasping is still a challenge, grasping real objects has been highly explored before, with researchers developing grasp taxonomies to provide a common terminology and framework of grasp patterns and how they change based on influencing factors such as object characteristics (Feix et al., 2014b) and task (Feix et al., 2014a) as emphasized in Chapter 3.

Taxonomies have also been highly used in HCI, due to the number of available technologies and interactions emerging, which leads to an overwhelming situation for researchers, users and developers who try to understand them from different viewpoints. Therefore, HCI researchers developed taxonomies to help them reason, compare, elicit and create the appropriate techniques for the problem at hand, in various domains such as gesture-based systems (Scoditti, Blanch, & Coutaz, 2011), User Interfaces (UI) (C. O. Seneler, Basoglu, & Daim, 2008), voice commands (Pérez-Quiñones, Capra, & Shao, 2003), data visualisation (Kleinman, Preetham, Teng, Bryant, & Seif El-Nasr, 2021), VR (Muhanna, 2015), and AR (Hertel et al., 2021).

This chapter presents existing literature on taxonomy development in HCI, reviewing methods for collecting and synthesizing data in taxonomies. Inspired by current HCI methods, VR limitations and state-of-the-art grasp taxonomies in real environments as presented in Chapter 3, a method for developing the first VR

Taxonomy of Grasp Types is proposed. The method proposes elicitation studies following Wizard of Oz methodology to be conducted for grasp data collection, and a labelling process inspired by real grasping literature to label and synthesize the grasps in a taxonomy. The baseline environment and labelling process conducted for every user experiment presented in this thesis are also detailed in this chapter.

This chapter is structured as follows: Section 4.2 presents taxonomies in HCI, Section 4.3 presents data collection methods for taxonomies, Section 4.4 presents the novel method for developing a VR grasp taxonomy, Section 4.5 presents the baseline environment for elicitation studies, Section 4.6 presents grasp classification metrics, Section 4.7 presents grasp labelling and Section 4.8 presents method overview where the modifications of the fundamental method for each user experiment presented in the thesis is detailed.

4.2 Taxonomies

Taxonomies have been used for a long time, from the work of Aristotle, who was the first to classify all living things, to taxonomies being developed for realising the metaverse today (S.-M. Park & Kim, 2022) and have highly contributed to technological advancements in various domains. This contribution is evident in the literature, with taxonomies developed decades ago still being addressed, improved and used as guidelines for emerging research work. In 1997, Gabbard (Gabbard, 1997) developed a taxonomy to lay a scientific foundation for developing high-impact methods for usability engineering of VEs. He collected data from literature and user interviews and synthesized the results in a thorough classification, enumeration and discussion of usability issues which has influenced emerging research work in VR (Papadimitriou, 2019). Later, Agah and Tanie

(Agah & Tanie, 1999) presented a taxonomy of research on human interactions with intelligent systems. They classified interactions based on research approach, application, system autonomy, interaction distance and interaction media following an extensive literature review. The taxonomy has been later used for updated reviews of interaction modes (Lazaro et al., 2021) or frameworks for human robot interactions (Beer, Rieth, Tran, & Cook, 2017).

With taxonomies showing promising results for defining common terminology and informing new research directions for the HCI community, the development of taxonomies has become common as a required step in filling a gap in the literature. For example, Hertel et al. (Hertel et al., 2021) identified a gap in understanding common interaction techniques for current immersive AR systems and therefore developed a taxonomy that focused on two dimensions: task and modality. For this, they conducted an extensive literature review and determined the characteristics of the taxonomies based on trends and keywords which were tagged iteratively by a group of researchers. The authors emphasized the importance of this taxonomy for novice researchers as well as experienced researchers to obtain comprehensive information about interaction techniques and insights about trends in immersive AR environments. Karam and Schraefel (Karam & Schraefel, 2005) focused on understanding trends in gesture-based computer interactions and conducted a literature review and a tagging session where they found four main categories: gesture styles, application domains and input and output technologies. This taxonomy was created with the aim of building on the knowledge gained across different domains and moved gesture-based interactions out of the research labs and into everyday computing applications and has served as guideline for hand gesture classification algorithms (Gadekallu et al., 2021). Moreover, since gesture-based systems are still a challenge, researchers also used it to create new iterations of gesture-based taxonomies that take into consideration current technological advancements (Carfi & Mastrogiovanni, 2021). Scoditti et al. (Scoditti et al., 2011) identified the lack of a common framework for comparative analysis of gesture-based interaction techniques based on accelerometers. To address this, they proposed a taxonomy that encapsulates gesture styles, applications domains, input and output technologies to provide a systematic structure for comparing and developing new gesture interaction techniques. Following an extensive literature review, which also included existing taxonomies, the authors organised the data in a controller vocabulary where each interaction type was classified without ambiguity. The taxonomy showed to be useful for designers, which used the framework to predict difficulties that users would encounter with different interaction techniques (Maranan et al., 2014).

This approach of developing taxonomies by collecting data through extensive literature reviews has been widely used in HCI (Klompmaker, Paelke, & Fischer, 2013). However, researchers emphasized that the choice of search terms and databases when conducting literature reviews can lead to incomplete taxonomies (Jonas et al., 2019). Moreover, Kreimeier and Gotzelmann (Kreimeier & Gotzelmann, 2020) showed that key points regarding the studied subjects are in some cases missing due to incomplete literature, therefore resulting in taxonomies that do not cover every aspect of interest. To address these limitations, researchers also looked into involving the user in collecting data for taxonomies, showing that additional user studies can further validate the usefulness and coverage of existing data collected during literature reviews (Adamides, Christou, Katsanos, Xenos, & Hadzilacos, 2015).

Consequently, aiming to provide an overview of important parameters for developing HCI interfaces, Seneler et al. (C. O. Seneler et al., 2008) created a taxonomy based on data collected directly from users, by conducting user interviews, focus groups and brainstorming sessions with experts, asking them strategic questions about technology adoption metrics such as task, product and information content in common HCI techniques. Due to its completeness, this work has been further used as a guideline in exploring technology adoption for various systems such as smart-phones (Aldhaban, Daim, & Harmon, 2015) or online services (C. Seneler, Daim, & Basoglu, 2010). A similar approach was followed by Perez-Quinones et al. (Pérez-Quiñones et al., 2003) who developed a taxonomy of voice interfaces by conducting several focus groups and interviews which later informed the development of voice-based multi-modal user interfaces (Schneider & Balci, 2006). While user interviews have shown to produce interesting insights for taxonomy development, researchers showed that for taxonomies of interaction techniques, observing the user while naturally interacting with a system might provide a better understanding of how the original experience changes for various HCI systems (Jain et al., 2021). Therefore, aiming to develop a taxonomy of activities that users engage in when interacting with game data visualisation, Kleinmann et al. (Kleinman et al., 2021) conducted a qualitative usability study where they asked participants to use the visualization system for a set of predefined tasks, to directly observe their approach for intuitive interactions. Other works focused on conducting elicitation studies to involve the users in the design phase instead of evaluating already existing systems. Yan et al. (Yan et al., 2018) conducted an elicitation study and invited participants to design physical gestures for a predefined set of objects in VR, which were then synthesized in a gesture taxonomy. Camp et al. (Camp, Schick, & Stiefelhagen, 2013) conducted a user study to observe the user while performing one-arm clicking gestures for distant interaction and developed a taxonomy for clicking gestures based on data collected from the participants. While there is a large body of work in taxonomies focused on literature reviews, involving the user in collecting data has shown promising results in achieving accurate and complete taxonomies, which increased the popularity of this method within the HCI community, with various interaction taxonomies being developed for understanding and improving interaction systems.

This section presented an overview of methods, use and contribution of taxonomies in the HCI community. Taxonomies are often built on top of other taxonomies or methods for classifying knowledge (Motejlek & Alpay, 2021) and therefore contributing to achieving deeper knowledge on specific topics and developing common terminology for smooth adoption and improvement of concepts by the community (Jain et al., 2021). While taxonomies were often used as a way of organizing literature reviews in a way that is accessible for other researchers and then serve as a guideline for further research work (Karam & Schraefel, 2005), researchers also focused on conducting user studies for collecting data when developing interaction taxonomies, to allow better observation and synthetization of common interaction types performed directly in the environment in question (Kleinman et al., 2021). This approach was also followed for developing grasp taxonomies in real environments, where grasp data was mainly collected through user observation (Feix et al., 2014a, 2014b). However, grasping interaction in VR has not been fully explored before (as presented in Chapter 2) and a grasping taxonomy for VR interactions has not been developed yet. The next section provides an overview of a novel methodology for developing grasping taxonomies in VR based on work conducted in both HCI and real grasping.

4.2.1 Taxonomy Development Methods

Taxonomies developed for real grasping were often using data collected from researcher's observation of their own hands and grasping actions (Napier, 1956) or collecting data by observing users grasp objects in their natural environment (M. R. Cutkosky, 1989; Kamakura et al., 1980; Feix et al., 2014a) as described in Chapter 3.

For interactions in virtual environments, researchers focused on conducting literature reviews (Kreimeier & Gotzelmann, 2020) or following research methodologies that involve the user in generating data included in the taxonomy such as user interviews (Jain et al., 2021), focus groups (C. O. Seneler et al., 2008), evaluative user studies (Kleinman et al., 2021), where the user is asked to use an existing system to evaluate its usability and limitations, or generative user studies (also referred to as elicitation studies), where the user is asked to design new interactions for a specific environment (Yan et al., 2018). However, a method for developing grasping taxonomies for VR has not been proposed yet.

To allow future development of intuitive grasp models in VR, a grasping taxonomy needs to provide common terminology and a framework for grasping patterns and their influencers, inspired by the work of (Feix et al., 2014a, 2014b) who analysed parameters that influence grasping in real environments, insights that were later used for developing and improving robotic grasp models (Elangovan et al., 2022). Therefore, the VR grasp taxonomy proposed in this thesis focuses on providing a framework for grasp patterns and parameters that influence grasping approaches in VR, that can be later used as decision-making tools where designers and researchers can search for existing guidelines on developing virtual grasp models. For this, a methodology for developing grasp taxonomies based on current literature is provided.

While methods for developing virtual grasp taxonomies have not been presented, a notable method for taxonomy development for virtual environments was proposed by Nickerson et al. (Nickerson, Varshney, & Muntermann, 2013) who emphasized the three important steps for developing a complete and accurate taxonomy: collection of the data, definition of taxonomy dimensions, and classification of the collected data in the dimensions proposed. Moreover, as part of their methodology the authors suggest that researchers should focus both on literature and case studies to examine the field and solve the situation of having limited empirical taxonomy data, and therefore suggest an iterative methodology. This iterative methodology refers to researchers starting with a subset of objects/parameters that they want to classify. Next, the researchers identify characteristics of these objects and group them in dimensions of the first iteration of the taxonomy. The first taxonomy is then reviewed and additional conceptualizations that might not have been identified in the first iteration are defined, and another dataset is collected and grouped in another taxonomy iteration. This process is repeated until the taxonomy has sufficient comprehensiveness and extendibility, however, this closure is subjective and difficult to define, with more work required to clarify it. This iterative methodology has been used for developing taxonomies in HCI (Hertel et al., 2021) and also furthered by Omair et al. (Omair & Alturki, 2020) to provide a framework of all forms of reasoning logic that are used in this process. They emphasize the important steps between data collection and account of findings which include: managing collected data, describing taxonomy dimensions, classifying data in dimensions and developing and assessing interpretations.

Other methods that have been explored for taxonomy development include the principles of faceted analysis approach (Kwasnik, 1999). The essence of facet analysis is sorting of terms in a given field of knowledge into homogenous, mutually exclusive facets, each derived from the parent universe by a single characteristic of division (Kwasnik, 1999), with the main challenge being to build classifications that are flexible and can accommodate new phenomena. Over time, facets have been reinterpreted and used in various domains including VR and AR taxonomies (Motejlek & Alpay, 2021). Another notable method for developing taxonomies for interaction techniques in VEs is the work of Bowman and Hodges (Bowman & Hodges, 1999) who proposed testbed evaluations, also known as experimental research to analyse interaction techniques for taxonomy development.

Their method works through creation of testbeds, environments and tasks that involve all of the important aspects of a task, that test each component of a technique, consider outside influences on performance and have multiple measures.

Taking into consideration the state-of-the-art methodology for developing grasp taxonomies in real environments (Feix et al., 2014a, 2014b) and methodologies for developing interaction taxonomies in VEs, a method for VR grasp taxonomy development is presented. Following the work of Feix et al. (Feix et al., 2014a, 2014b) in real environments and work in virtual environments (Nickerson et al., 2013; Omair & Alturki, 2020) the method for developing a VR grasping taxonomies follows the main stages of taxonomy development: data collection, definition of taxonomy dimensions and classification.

When collecting data for HCI taxonomies, researchers focused on literature reviews and user experiments. Since grasping interaction has not been fully explored in VR, literature reviews for collecting taxonomy data would not provide the required knowledge and therefore data collection in this thesis is done through user experiments. When conducting user experiments for real grasping analysis, Feix et al. (Feix et al., 2014a, 2014b) observed two machinists and two housekeepers while using their hands during normal professional work, for at least eight hours per subject. However, due to current nature of VR technology, observing users while interacting with objects in VR for long periods of time could lead to cyber-sickness (Martirosov, Bures, & Zítka, 2022) and therefore the use of testbed environments is preferred (Nieuwenhuizen, 2015; Mayer et al., 2021). Hence, this thesis presents a set of experimental research studies to analyse grasping interaction patterns in VR and how these are influenced by different parameters. The methodology for collecting data through user experiments is presented in more detail in Section 4.3.

For defining the taxonomy dimensions, HCI taxonomies focused on relevant categories for providing a complete overview of the metrics evaluated (Jain et al., 2021). To ensure a complete analysis of grasping patterns in VR, the work presented in this thesis followed the methodology of Feix et al. (Feix et al., 2014a, 2014b) who identified the main parameters that influence grasping in real environments (object characteristics and tasks) and the metrics that commonly describe a stable grasp posture (grasp category and grasp type) as presented in Chapter 3. To facilitate the development of a complete VR grasp taxonomy, this work followed an iterative methodology (Nickerson et al., 2013), meaning that the taxonomy presented in this thesis is developed in iterations, over multiple user studies.

To classify the collected data, this work followed the methodology of Feix et al. (Feix et al., 2014a, 2014b) where the grasps were labelled by researchers using the common terminology of grasp types (Feix et al., 2009). This method of tagging/labelling collected data by experts against a predefined framework is common in HCI taxonomy development (Hertel et al., 2021). The methodology for labelling and classifying the data is presented in more detail in Section 4.7.

4.3 Data Collection for Taxonomies

For selecting the most appropriate methodology for collecting data directly from the user in a testbed environment, research methods used in HCI for observing user behaviour in a testbed environment were reviewed. Evaluative research methods refer to methods that involve the user in testing existing prototypes, products or interfaces by collecting performance measures such as task speed, accuracy and preference measures such as aesthetic and emotional response (Kleinman et al., 2021). These methods are often used to evaluate usability of VR environments and have been highly used to assess problems and limitations of interaction in

VR (Paes & Irizarry, 2018). While these methods are highly used and provide insightful results for interaction designers, they can only be used on prototypes or products that have already been developed and need improvement. Since grasping in VR has not been fully explored before and a natural and intuitive grasping model does not exist, evaluative research methodologies are not appropriate to use for the experiments conducted in this thesis. This emphasized the need for evaluating design research methods, where the user informs design and development of new prototypes and systems based on user feedback that can be classified in taxonomies (Camp et al., 2013). Various design research methods have been explored in HCI to inform design decisions, including user interviews, surveys, diary studies, contextual inquiry, experience sampling and other mixed methods, however, for developing interaction taxonomies researchers focused on asking the users to propose gesture sets that are natural and intuitive for various tasks and objects. This method is known as elicitation and has offered significant contributions towards the design of intuitive user interfaces (Ortega et al., 2019) and interaction taxonomies (Wobbrock, Morris, & Wilson, 2009; Rodriguez & Marquardt, 2017; Ali, Morris, & Wobbrock, 2021).

4.3.1 Elicitation

Gesture elicitation is a technique that emerges from the field of participatory design (Morris et al., 2014) which aims to enable those who will use the technology to have a voice in its design, without needing to speak the language of professional technology design (Simonsen, 2012). Elicitation studies refer to a methodology in which the experimenter shows a referent, known as the effect of an action or features of the user interface that can be controlled independently using a command, and asks participants to perform the interaction that would produce that effect (Villarreal-Narvaez, Vanderdonckt, Vatavu, & Wobbrock, 2020). As emphasized in Chapter 2, a big challenge in HCI is finding a feasible interaction dictionary that is easy to remember and intuitively performed by users. To address this, researchers started to use gesture elicitation studies, which are now popular and resourceful for informing the design of intuitive gesture commands, reflective of end-users' behaviour, for controlling all kinds of interactive devices, applications and systems (Villarreal-Narvaez et al., 2020). The approach of prompting users with referents (also known as effects of an action such as a button that elicits a click) and having them perform signs or causes of those actions dates back to 1984 when Good et al. (Good, Whiteside, Wixon, & Jones, 1984) used this approach to develop an intuitive command-line e-mail interface. Later, Nielsen et al. (Nielsen, Stoerring, Moeslund, & Granum, 2003) proposed a procedure for designing a gesture vocabulary for hands-free computer interaction in ubiquitous computing, focusing on taking into account users' viewpoint regarding intuitive interactions, ergonomics and learning rates. The procedure focuses on identifying the functions required by the application and collecting the gestures from the user domain by taking users through scenarios where they are required to communicate the functions previously identified, with this communication being recorded for further analysis.

Subsequently, Wobbrock et al. (Wobbrock, Aung, Rothrock, & Myers, 2005; Wobbrock et al., 2009) developed the end-user elicitation study method, aiming to make interactive systems more guessable, learnable and usable. First, they introduced the user elicitation approach for symbolic input in the form of a guessability method, showing that high guessability is desired for successfully meeting users' attempts at performing gestures, typing commands or using buttons despite their lack of knowledge of the relevant symbols (Wobbrock et al., 2005). For this, they asked participants to propose symbols for specified referents and introduced a measure for agreement among symbols proposed by participants. This method showed to increase symbolic input guessability of the unistroke alphabet EdgeWrite (Wobbrock, Myers, & Kembel, 2003). Further, Wobbrock et al. (Wobbrock et al., 2009) used this approach for designing tabletop gestures that rely on eliciting gestures from non-technical users and synthesized the results in a user-defined taxonomy of surface gestures.

Over time, elicitation studies have been highly used for developing user-defined taxonomies in HCI, with most of the research surrounding freehand input justifying the chosen interaction paradigms based on elicitation studies (Piumsomboon et al., 2013). Chen et al. (T. Chen, Xu, Xu, & Zhu, 2021) used an elicitation method to develop a gesture-based interaction technique that leverages stereo microphones in a commodity smart-phone to detect the tapping and the scratching gestures on the front, the left and the right surfaces on a mobile VR headset. The referents for the elicitation were displayed on a monitor in front of the participant with the animation playing the effects of actions. Based on these, participants were asked to design the gestures for the referents which were then synthesized in a taxonomy containing 150 user-defined gestures. Angelini et al. (Angelini et al., 2014) focused on gesture interaction as a modality for reducing driver distraction and increase safety while driving. For this, they asked 40 participants to elicit six gestures during a driving simulation, with no predefined commands assigned. Based on the results of this experiment, they developed a taxonomy of gestures performed on a steering wheel. This approach allowed users to intuitively interact with the system in a naive manner, facilitating the design of natural interactions that are easy to use and remember. However, the authors noticed different behaviours among users; some participants were concerned about safety and were looking for gestures to perform without releasing the hand from the steering wheel while other participants were looking for very simple gestures and were not concerned by the safety issue. This is due to the fact that users elicited gestures in front of a steering wheel while simulating driving but not while in a real driving experience, which could have lowered the awareness that some gestures could be dangerous to be performed during the drive, as many of them require leaving one or both hands from the steering wheel. This suggests that elicitation studies are more likely to produce realistic insights for interaction design when they are conducted in the same or similar environment in which the interaction paradigms will be used. This method can therefore successfully substitute observation experiments that are performed for real grasping taxonomies (Feix et al., 2014a, 2014b) by immersing the users in a VR environment where they are asked to propose grasps for a variety of referents.

Grasping elicitation for virtual objects has been conducted by Yan et al. (Yan et al., 2018) who asked users to design grasping gestures for a list of objects. They told each participant the name of an object and asked them to recall the hand gestures they would normally do to grasp that object in reality. They synthesized their results in a gesture taxonomy and provided insightful results in terms of orientation of the hand, position and posture. These gestures were then transferred to VR, however they showed that users still had to learn the gesture mapping to successfully recall the grasping gestures. Therefore, using virtual referents when eliciting grasping patterns in VR instead of collecting information that is entirely based on user recollection of grasping actions in reality, might provide clearer insights into users' intuitive grasping approach. Yet, this approach of asking users verbally to describe and motivate gestures has been widely used in HCI, and is also referred to as think-aloud protocol (Ericsson & Simon, 1984) where participants verbally describe their thinking process while being presented with referents.

Sharma et al. (Sharma, Roo, & Steimle, 2019) conducted an elicitation study following the think-aloud protocol for developing a systematic understanding of the complex relationship between micro-gestures and various types of grasps. Their aim was to develop a taxonomy of micro-gestures that are performed by users while holding another real object to be used for developing computational methods that work in settings where the user's hands are busy holding an object. However, they intentionally refrained from using any sensing technology so as to not bias the user's response by restrictions imposed by equipping everyday objects with sensors and therefore no interaction feedback was provided. A similar approach was followed by Oh and Findlater (Oh & Findlater, 2013) who conducted a study to understand the end-user gesture creation process. For this, participants completed a single one-hour session consisting of: open-ended gesture creation, gesture creation for specific actions and judging feature saliency. They used the think-aloud methodology and did not provide interaction feedback to participants for the first two tasks. In the third task, the authors used a simple gesture recogniser to analyse recognition accuracy of the newly proposed gestures. They found that participants often accounted for the perceived abilities of the recognition system, although their understanding of the system was not always accurate. This led to lower recognition rates of the newly proposed gestures, showing that the lack of interaction feedback in this method might influence the intuitive interactions provided by users.

Another limitation of elicitation studies where users do not have any interaction feedback was emphasized by Wobbrock et al. (Wobbrock et al., 2009) who noted that application context could influence user behaviour and therefore their choice of gestures and Cooke and Cuddihy (Cooke & Cuddihy, 2005) who noted that users may not be especially proficient at verbalizing their thoughts, even after they have been trained in speaking as they concurrently perform tasks. A notable methodology used in HCI elicitation studies that aimed to facilitate interaction feedback is the Wizard-of-Oz methodology which can be combined with elicitation.

4.3.1.1 Wizard of Oz

A *Wizard of Oz* methodology is defined as the experiment approach where the interaction is mediated by a human operator to allow the user more freedom of expression or constrain the interaction in a systematic way (Dahlbäck, Jönsson, & Ahrenberg, 1993). Kelley (Kelley, 1985) first coined the phrases "Wizard of Oz" and "Oz Paradigm" for this purpose in 1980 to describe the method he developed where he used a blackout curtain separating him as the experimenter ("wizard") from view by the participant during the study. Elicitation studies typically involve a Wizard of Oz approach in which gestures are elicited from users by first portraying the effect of a gesture as demonstrated by an unseen technical wizard manipulating the system and then asking the users to perform its cause (Wobbrock et al., 2009).

The Wizard of Oz methodology has proven to be an important tool for collecting data directly from the user that can then be synthesized in interaction taxonomies, highly contributing to developing complex interactive applications in VR/AR. For collecting information on how users intuitively interact with a system, Nielsen et al. (Nielsen et al., 2003) emphasize the importance of taking the testees away from technical thinking especially when conducting tests on technically minded people, to avoid them thinking in terms of interfaces and algorithms. This can be achieved using the Wizard of Oz methodology and has been highly used as part of interaction design processes (K. Kim, Park, & Lim, 2021), especially for designing hand interactions (M. Lee & Billinghurst, 2008; Alce, Wallergard, & Hermodsson, 2015). Due to the benefits it proposes, researchers highly integrated this method in VR/AR research: Alce et al. (Alce et al., 2015) developed WozARd, a Wizard of Oz method created for wearable AR interactions. This method lets the user interact with the system through a smart-watch allowing the human operator to easily change the UI without reprogramming the application. Belluci et al. (Bellucci, Zarraonandia, Díaz, & Aedo, 2021) presented Welicit, a Wizard of Oz tool to support researchers in running VR elicitation studies. The system provides tools to adapt the elicitation methodology to immersive environments by allowing users to experience the result of the proposed interactions. Speicher et al. (Speicher, Lewis, & Nebeling, 2021) presented a new method and tool for rapid prototyping of AR/VR experiences and implemented dedicated support for Wizard of Oz to allow simulations of spatial interactions. Williams et al. (A. S. Williams, Garcia, & Ortega, 2020) used the Wizard of Oz for understanding how people naturally manipulate virtual objects in AR using multi-modal interactions such as gesture and speech. This is done by observing participants perform these interactions in a natural and intuitive manner in the VR environment, as per Wizard of Oz's methodology.

4.4 Taxonomy Method for VR Grasping

Taxonomies for real object grasping have been generally developed by collecting grasp data through observation, either observing the researcher's own hands (Napier, 1956) or observing humans during their professional work (M. R. Cutkosky, 1989) and synthesizing the results in a grasp taxonomy that categorises grasps based on relevant grasp metrics. However, replicating this methodology in VR is challenging due to technological limitations (such as wearing a HMD for long hours which might lead to cyber-sickness). Observing the users during their professional work for eight hours as in the work of Feix et al. (Feix et al., 2014b, 2014a) is therefore impossible and the methodology needs to be adapted to track user behaviour for shorter periods of time.

To evaluate user behaviour in virtual environments for specific tasks but a shorter period of time, work in HCI has proposed creating testbed evaluations (Bowman & Hodges, 1999) where users are immersed in environments that involve all important aspects of a task and test each component of a technique that consider influencing parameters on performance. The work in this thesis follows this approach and presents virtual environments where the user is immersed and asked to intuitively interact with the environment. For each user experiment, the environment is designed to present parameters that might influence grasping patterns in VR, carefully selected through literature reviews. Considering the gap in VR grasping literature, a design methodology needs to be followed to collect data in the absence of a grasping system and inform design and development of future intuitive grasp models. For this, HCI researchers focused on elicitation methods where users are directly involved in designing hand interactions for predefined referents. In this novel methodology for VR taxonomy development, the referents are the influencing parameters (such as virtual objects and tasks) and the users propose grasps for each of these referents, which are then synthesized in the first VR grasping taxonomy.

Another limitation in replicating grasping observation studies in VR is that virtual objects do not react automatically in the absence of a grasping algorithm that would constrain user grasp choice to grasping patterns that are recognised by the system. For this, a Wizard-of-Oz methodology is chosen, where interaction is mediated by a human operator to allow more freedom of expression. This way, every time the user grasps an object in the testbed environment, the human operator triggers the interaction between the hand and the virtual object, regardless of the grasp pattern performed. Similar to work in real grasping, a camera is attached on the users' head, together with the HMD and hand tracking sensor to record their hands at all times (Huang, Ma, Ma, & Kitani, 2015). This way, users' intuitive and natural behaviour towards grasping in VR is collected for different influencing parameters to facilitate the development of a complete VR grasping taxonomy. This elicitation together with Wizard-of-Oz method was applied for all VR user experiments presented in this thesis.

4.4.1 Limitations

While elicitation studies are critical to understand gestures and preferences from users, this practice has also generated debate and contradiction from the community. One important limitation of elicitation methods is legacy bias, which is a form of bias introduced into interaction proposals by participant familiarity with prior interaction techniques and technologies. Morris et al. (Morris et al., 2014) shows that users propose legacy-inspired interactions for several reasons: an explicit desire to transfer their knowledge of past systems to new ones, a desire to minimize physical and mental exertion when interacting in new modalities, and misunderstandings of the fundamental capabilities of novel sensing technologies. Such biases may cause gesture elicitation methods to get caught in local minima, failing to uncover interactions that may be better suited for a given medium than those that leap rapidly to users' minds. Reports from research studies employing gesture elicitation note many examples of legacy bias. For example, Wobbrock et al. (Wobbrock et al., 2005) noted that, despite presenting participants with a large multitouch touchscreen without UI elements from traditional PC interfaces, most participants suggested mouse-like single-point or simple-path gestures. In a multi-modal gesture and speech elicitation study, Morris et al. (Morris, Wobbrock, & Wilson, 2010) showed a similar finding with one participant referring to his hand as the "mouse" or another participant who avoided bimanual interactions because they might require a system with two cursors. Similarly, Oh and Findlater (Oh & Findlater, 2013) found that participants tend to focus on the familiar, even when instructed to create novel gestures.

Several methods have been proposed to mitigate legacy bias (Morris et al., 2014) such as: Priming, which is directly or indirectly influencing participants' mind-set before running a study (for example description of the types of gestures being elicited) and Production, which is asking participants to produce N+ (usually three or more) proposals per referent. Production is based on the idea that participants will exhaust their legacy proposals if asked for more than one proposal per referent. This would cause later proposals to be less biased; Hoff et al. (Hoff, Hornecker, & Bertel, 2016) adapted and experimentally tested Morris et al. (Morris et al., 2014) 's suggestions for reducing legacy bias and found that the practical effectiveness of these strategies might be limited, given the fact that they only found medium effect sizes and a wide variance between participants that overshadows any effects. Priming resulted in less legacy gestures and a quicker generation of ideas, but the difference was not of statistical significance. However, they state that it is possible that letting participants suggest more than three gestures might reduce legacy gestures. Partnered elicitation is another proposed solution that refers to grouping participants in pairs during the elicitation study. When using paired elicitation, participants are placed into small groups of three people. The goal is to cause increased variety in the elicited proposals by using collaboration. There were a few studies using partnered elicitation (McAweeney, Zhang, & Nebeling, 2018), however, the effect on reduction of legacy bias is unknown.

The methodology presented in this thesis takes into consideration these limitations with priming and production being implemented individually for each user experiment. Partnered elicitation was not suitable for the presented experiments, as it also requires the implementation of the think-aloud protocol, which has shown to present some limitations in terms of user communication and the focus of this work was to allow grasp interaction directly in VR in an intuitive way. Asking users to collaborate to come up with a grasp for each referent would therefore break the presence in VR and might lead to changed grasping behaviour. The protocol for each experiment along with priming and production methods is presented in the following chapters. However, if these measures did not mitigate legacy bias, research has also showed benefits linked to this limit. As participants draw upon culturally shared metaphors, they tend to propose similar legacy inspired interactions, resulting in high agreement scores in elicitation studies. This agreement indicates that legacy-inspired interactions are easily guessable and learnable and perhaps appropriate for systems intended to provide natural and intuitive interactions (Morris et al., 2014).

While Wizard-of-Oz methodology brings numerous advantages to interaction design such as instant user feedback or informed design based on user behaviour, there are also limitations associated with this method. Nebeling et al. (Nebeling, Huber, Ott, & Norrie, 2014) showed that using Wizard-of-Oz for eliciting multimodal interactions might not be suitable to obtain accurate interaction proposals, showing that users choose different interaction patterns for the same referent based on the feedback received during the experiment. The work presented in this thesis focuses on uni-modal interaction only (only looking for hand input) and user experiments are conducted by the same "wizard" to provide the same feedback for referents and not bias the interaction choice (Nebeling et al., 2014).

Technique	Relevance to this work
Taxonomy development method for VE (Nickerson et al., 2013)	The main steps for developing a taxonomy were considered in this work
Iterative taxonomy development (Hertel et al., 2021)	An iterative process was used in this work.
Faceted analysis (Kwasnik, 1999)	The taxonomy presented in this work took this into consideration and structured grasps based on multiple attributes.
Testbed evaluations (Bowman & Hodges, 1999)	This work focused on creating tasks relevant for analysing virtual grasping.
Literature reviews (Klompmaker et al., 2013)	This work presents a literature review for grasping and HCI literature to inform the method and structure of the VR taxonomy.

Table 4.1: Overview of methods used for developing taxonomies in HCI. For each method, the key aim is presented and the relevance to this work is described.

4.5 Baseline Environment

This section presents a concise description of the environment developed for conducting the elicitation studies presented in this thesis.

4.5.1 Apparatus

The baseline environment used the Oculus Rift DK2 VR Head Mounted Display (HMD) (FOV = 100°) with the Leap Motion (FOV = 135°) tracking device mounted on the front of the HMD, facing the participants' hands to track hand movements during the interaction. The Leap Motion Controller is a standard interface controller, which is becoming common within the VR devices for freehand tracking and fine manipulations (Guzsvinecz et al., 2019).

The sensor has a depth between 10 cm to 60 cm and a field of view of 140°x 120°. The LEDs in the controller illuminate the hands with infrared light and pulse in sync with the camera framerate, sending data back to the computer to track the hands. Then, the software uses images to generate a virtual model of hand movements including not only palm and fingertips, but the joints and bones inside the hand. Once the image data is streamed to the computer, the Leap Motion Controller applies advanced algorithms to the raw sensor data and processes the images through the Leap Motion Service and reconstructs a 3D representation of the hand (Ultraleap, 2021). Using the Leap Motion Controller together with the Oculus Rift is a popular configuration for VR applications such as medical training (Fong et al., 2021; Augestein, Kortemeyer, & L, 2021), simulations (Dhanasree, Nisha, & Jayakrishnan, 2018) and robotics (Hameed et al., 2017; Yang, Chen, & Zhu, 2020).

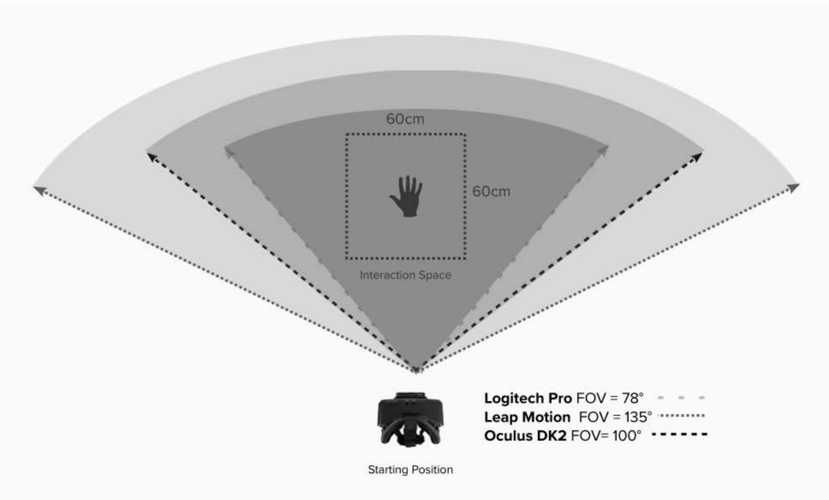
To record participants' real hand during grasping interaction, a Logitech Pro 1080p

HD (FOV = 78°) webcam was positioned on top of the Oculus DK2 (Figure 4.1 b). As shown in Figure 4.1, the virtual interaction space is constrained by Leap Motion Controller's FOV and is 60 cm $\times$ 60 cm $\times$ 60 cm. Participants are seated in the experimental environment, which is a controlled environment under laboratory conditions, lit by a 2700k (*warm white*) fluorescent with no external light source. Across all the studies presented in this thesis, the physical configuration of the system strictly and consistently followed the recommendations of the Oculus DK2 and Leap Motion Controller's manufacturers to ensure ideal operating conditions of the sensor. Grasp data was measured from the Leap Motion Controller and the Logitech webcam. This baseline environment is used for all VR experiments presented in this thesis.

4.5.2 System Architecture

The VR environment used for the elicitation studies presented in this thesis was developed in a controlled laboratory environment using the following tools:

- **Unity 2018.2** Unity is a cross-platform game engine developed by Unity Technologies that supports a variety of desktop, mobile, console and virtual reality platforms. The engine can be used to create 3D and 2D games as well as interactive simulations. The engine offers a primary scripting API in C# and allows both creating and importing predefined 3D models. It supports building, testing and publishing VR applications and was used for creating the virtual environment, importing the 3D virtual models, recognising and showing hand movements in VR, triggering interactions and collecting grasping data.
- **Oculus Integration Package** Oculus Integration Package is a Unity plug-in that allows building applications for the Oculus with the correct framework.



(a) Configuration

(b) Apparatus

Figure 4.1: System configuration displaying the custom experimental framework: Leap Motion and Logitech Pro 1080p HD camera attached to the Oculus Rift DK2.

Moreover, it contains Oculus VR utilities, a set of scripts and prefabs to enable VR development and was integrated with Unity for development of the VR environments presented in the experiments of this thesis.

- **Leap Motion 4.0 SDK** Leap Motion SDK is a collection of software development tools that facilitate the use of Leap Motion Service within the VR environments presented in the experiments of thesis. The SDK contains two basic libraries that define the API to the Leap Motion tracking data. The two libraries (C and C++) contain wrapper classes that define language bindings for C#, which is the programming language used for developing the VR environments presented in this thesis.

- **Autodesk Maya** [1]: Maya is a 3D computer graphics application and is used for modelling 3D virtual objects, namely cubes that were used for the training session in each experiment presented in this thesis. Additionally, Maya allows properties of an existing 3D object such as complexity, texture, size and shape to be adjusted which was useful for developing the virtual environment and achieving the required visual quality of the virtual objects. Maya allows export of ".OBJ" files that can be directly imported in the Unity scene.

4.5.3 Virtual Objects

Previous user elicitation studies have used pictorial (Wobbrock et al., 2009; Wittorf & Jakobsen, 2016) or animated (Piumsomboon et al., 2013) referents to encourage participants to develop their own set of gestures based on showing the effects (*referents*) these will have on the system. The work in this thesis aims to explore natural grasping patterns in realistic virtual scenarios therefore, referents

[1] http://www.autodesk.com/products/maya/overview

(a) Food items (b) Kitchen items (c) Tool items

(d) Shape items (e) Toys items (f) Toys - lego

Figure 4.2: Yale-Carnegie Mellon University-Berkeley Object and Model Set (Calli et al., 2015) providing graspable objects that are frequently used in daily life, designed to be used for grasping manipulation research and covering a variety of shapes, sizes and textures.

used in these experiments are 3D virtual representations of real objects.

When selecting an object set for assessing grasping interaction, various object attributes need to be considered as it has been shown they influence grasping interaction in real environments; Napier (Napier, 1956) showed that object size highly influences grasping pattern for real objects. Cesari and Newell (Cesari & Newell, 2002) showed that fewer fingers are used for grasping smaller objects as opposed to larger objects that require more fingers engaged in the grasp pose. However, studies to investigate how virtual object attributes influence grasping patterns in VR have not been conducted yet. To evaluate this and develop a detailed overview of how different objects are grasped in VR, a combination of objects of different shapes and sizes were used to elicit grasping interaction.

Object and model sets are generally the fundamental elements involved in benchmarks for manipulation, with substantial effort being put in providing databases of graspable objects to be used for grasping manipulation research (Singh, Sha, Narayan, Achim, & Abbeel, 2014; B. Li et al., 2014; Goldfeder, Ciocarlie, Dang, & Allen, 2009). Calli et al. (Calli et al., 2015) present the *Yale-Carnegie Mellon University-Berkeley Object and Model Set* which is designed to cover various aspects of the manipulation problem such as variety of shapes, sizes, and textures. The object set includes objects that are frequently used in daily life and considers objects that are used in simulations and experiments. The authors provide high-resolution RGBD scans, physical properties, and geometric models of these objects for easy incorporation in various software platforms. The objects in the *Yale-Carnegie Mellon University-Berkeley Object and Model Set* are divided into the following categories: food items, kitchen items, tool items, shape items and task items as shown in Figure 4.2. From this dataset, 16 virtual objects were chosen to cover a variety of shapes and sizes from each category. From food category the following objects were chosen: mustard, cracker box, meat can, gelatine box, orange and banana; From kitchen items the following objects were chosen: spoon, mug and cleanser bottle; From tool items the following items were chosen: clamp, hammer and scissors. From shape items only brick was chosen, since there was already a spherical object from food category (orange). From toys items a small Lego piece was chosen. These 16 objects chosen covered a wide range of shapes (cuboid, spherical, cylindrical, with/without handles and irregular shapes) and sizes (ranging from 14 mm to 250 mm). Table 4.2 presents the chosen objects with their size on each axis (X, Y and Z).

4.5.4 Tasks

Task intention has shown to influence grasping when approaching real objects (Feix et al., 2014b). Moreover, VR systems where natural interactions have an

Object Name	Object	x(mm)	y(mm)	z(mm)	Object Name	Object	x(mm)	y(mm)	z(mm)
Banana		36	36	190	Bleach Cleanser		250	98	65
Brick		50	75	50	Clamp		90	115	27
Cracker Box		60	158	210	Gelatine Box		28	85	73
Hammer		24	32	135	Lego		50	50	50
Marker		18	18	121	Meat Can		50	97	82
Mug		80	80	82	Mustard Bottle		58	95	190
Orange		73	73	73	Sponge		72	114	14
Spoon		14	20	195	Scissors		87	200	14

Table 4.2: Virtual object set selected from Yale-Carnegie Mellon University-Berkeley Object and Model Set (Calli et al., 2015). This table shows the name of each object, used as reference for the remaining of this thesis, visual representation of each object and X,Y and Z dimensions in mm. The virtual objects selected are: Banana, Bleach Cleanser, Brick, Cracker Box, Gelatine Box, Hammer, Lego, Marker, Mug, Mustard, Orange, Sponge, Spoon and Scissors.

impact on overall performance are task-oriented (Ma, Varley, Shark, & Richards, 2010). When manipulating virtual 3D content, translation, orientation and docking (6 DOF) are fundamental tasks for which the choice of interaction gestures is critical for usability and performance (Martinet, Casiez, & Grisoni, 2010). Therefore, to understand the pattern of grasp choice in VR, each user study presented in this thesis asked users to perform one fundamental task, either simple translation (one or more directions) or translation and rotation (docking) tasks (on different rotation angles). An overview of the tasks employed in each chapter is presented in Section 4.9 and detailed in the methodology sections of each chapter.

4.5.5 Virtual Environment

The VR elicitation studies were conducted in a controlled environment under laboratory conditions. The test room was lit by a 2700k (*warm white*) fluorescent with no external light source. The virtual environment was composed of a virtual table, the virtual objects and other specific virtual props as described in the following chapters. The virtual table was 1000 mm long and 600 mm wide. For each task evaluated in the thesis, the target location (position and rotation) was indicated by a replica of the virtual object, rendered in a green transparent texture. This replica indicated that subjects had to move the virtual object from the original position to this target position as shown in Figure 4.3.

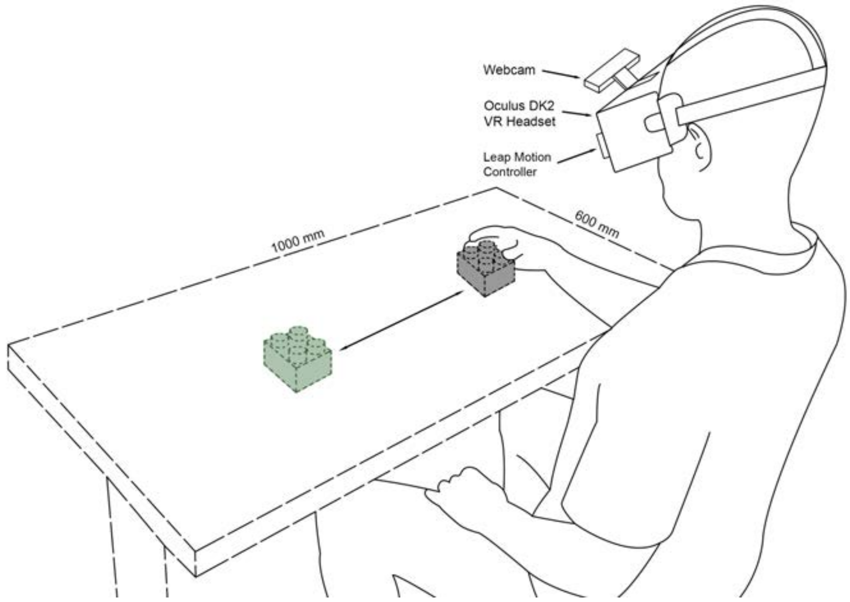


Figure 4.3: Virtual environment showing the interaction space, virtual object and target object.

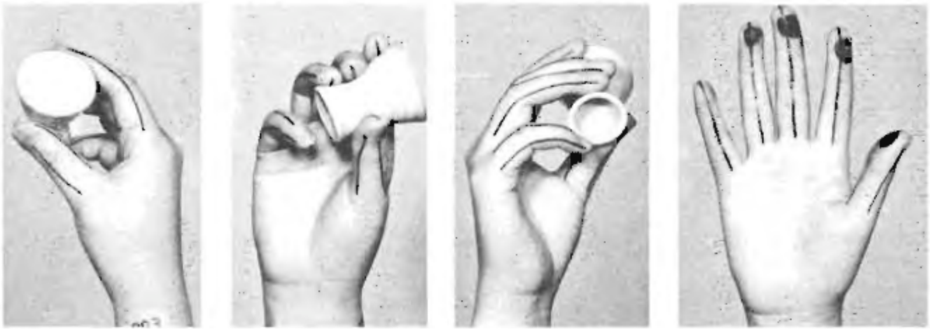


Figure 4.4: Methodology for grasp observation in real grasping taxonomy development as presented by (Kamakura et al., 1980)

4.5.6 Grasp Data

For collecting and storing grasp postures for real grasping taxonomy development Kamakura et al. (Kamakura et al., 1980) smeared the objects used in the study with ink and then asked participants to grasp them, which was followed by photographing the hand and analysing the contact areas of each grasp as seen in Figure 4.4. More modern approaches such as the work of Feix et al. (Feix et al., 2014b, 2014a) used a camera attached to users' heads to record the hands during eight hours of interactions with real objects which provided information both about hand posture and contact areas when touching the objects. Aiming to follow this real grasping taxonomy methodology for developing the first VR grasping taxonomy, both data about the posture of the hand and the contact areas when the hand interacts with the virtual objects is required for a detailed analysis of grasping patterns. While wearing the VR HMD, the user only sees the virtual environment, without seeing anything in the real environment. This means that if the data is captured using a virtual camera inside the VE, the recordings will show the virtual representation of the hand interacting with the virtual objects. To support more accuracy in the labelling process, the real hand was also recorded during the interaction following methodologies of real object grasping analysis

(Feix et al., 2014b).

Therefore, an additional camera is used to record users' real hand during the interaction, which was attached to users' head (on top of the VR HMD). This is needed on top of the recording of the virtual environment during grasping, as recording only the real hand would not provide information about contact areas between hand and object during grasping. Therefore, the grasp data collected during the elicitation studies presented in this thesis includes:

- **Grasp real view**: Using a web camera, images with the real hand during grasping interactions are captured and referred to as "real view" in this work (see Figure 4.5 a).
- **Grasp virtual view**: Using a virtual camera in the virtual environment, images of the virtual interaction (virtual hand and virtual objects) are captured during the grasping interaction and are referred to as "virtual view" in this work (see Figure 4.5 b). Recording both virtual and real view ensures a more accurate labelling of the grasp, as both the real hand pose and its position regarding the virtual objects are available to the rater. The labelling process is described in more detail in Section 4.7.
- **Hand Position**: The centre position (X,Y,Z) of the palm and fingers in millimetres from the hand tracking device.

4.6 Grasp Classification Metrics

Using the grasp data presented above, grasp aperture (GAp) was calculated using hand and finger position data and the grasps were classified in grasp categories, grasp types and grasped dimension using the real and virtual view.

(a) Real view (b) Virtual view

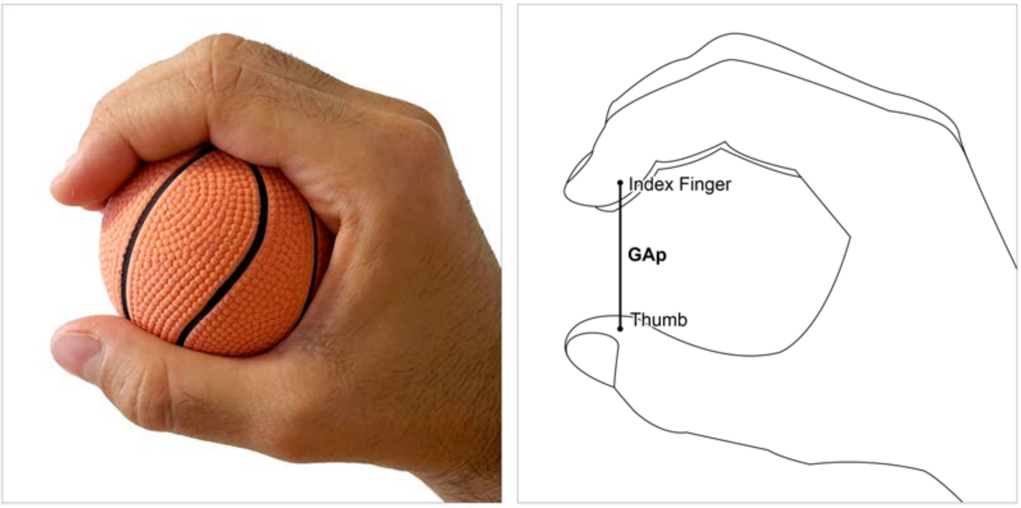
Figure 4.5: Grasp captures recorded during the user elicitation studies: a) shows an example of a real view image and b) shows an example of a virtual view image.

4.6.1 Grasp Aperture

Grasp aperture is a common metric in human manipulation studies (Al-Kalbani et al., 2016a) to measure how accurately users estimate the size of the virtual object. Grasp aperture (GAp) is defined in equation 4.1 to be the distance between the tip and index fingertip (Edsinger & Kemp, 2007):

$$GAp = \sqrt{(P_x - B_x)^2 + (P_y - B_y)^2 + (P_z - B_z)^2} \tag{4.1}$$

Where GAp is the distance between the index and the thumb fingers in the x, y and z axes, and P_x, P_y and P_z are the co-ordinates of the index fingertip, and B_x, B_y and B_z are co-ordinates of the thumb tip. A visual representation of this is shown in Figure 4.6.



(a) Power Sphere grasp example (b) GAp of the Power Sphere grasp

Figure 4.6: Grasp aperture (GAp) used for quantifying grasp accuracy

4.6.2 Grasp Labels

Following the work of Feix et al. (Feix et al., 2009), the work presented in this thesis uses the Human GRASP Taxonomy terminology, as it is the most complete taxonomy of grasp types to date. The taxonomy presents the three main grasp categories (Power, Intermediate and Precision) with grasp types within each category sub-categorised in Thumb Abducted and Thumb Adducted as described in Chapter 3. For easy reporting throughout the thesis, each grasp type was assigned a grasp code as shown in Figures 4.7, 4.8 and 4.9: Grasp types in power category were marked with a [P] and an index number, grasp types in intermediate category were marked with an [I] and an index number and grasp types in precision category were marked with a [PC] and an index number.

Power category presents the following grasp types: Large Diameter [P1], Small Diameter [P2], Medium Wrap [P3], Ring [P4], Power Disk [P5], Power Sphere [P6], Sphere 4 Finger [P7], Sphere 3 Finger [P8], Distal Type [P9], Extension

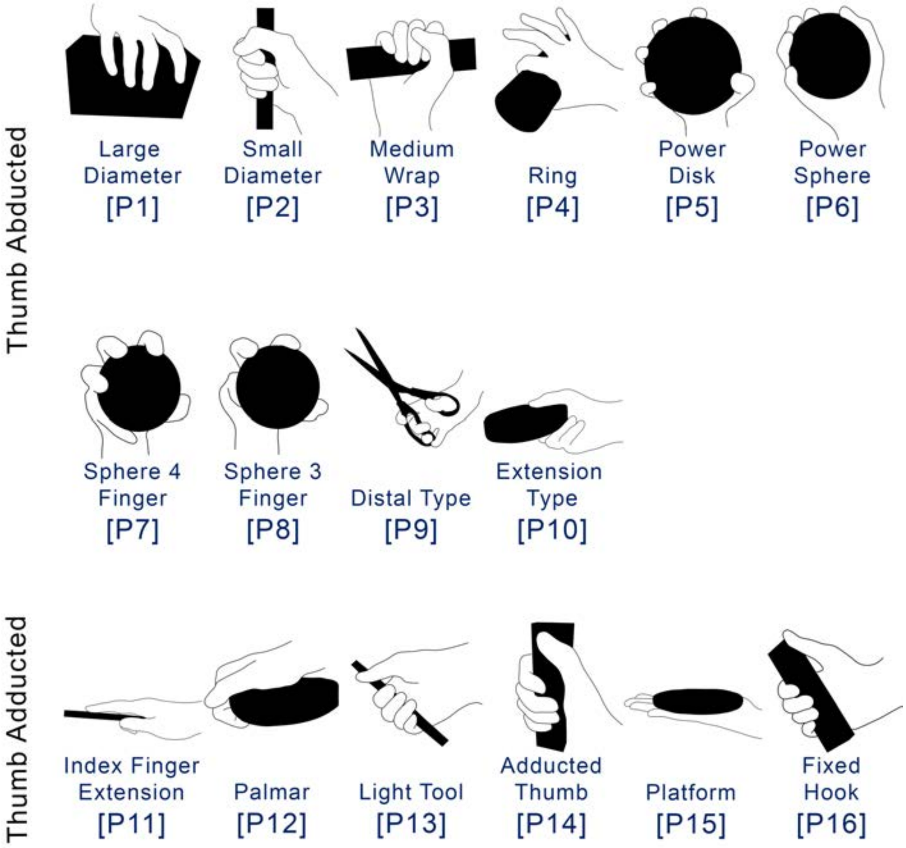


Figure 4.7: Power Grasps from the Human GRASP Taxonomy (Feix et al., 2009)

Type [P10], Index Finger Extension [P11], Palmar [P12], Light Tool [P13], Adducted Thumb [P14], Platform [P15] and Fixed Hook [P16].

Intermediate category presents the following grasp types: Adduction [I1], Tripod Variation [I2], Lateral Pinch [I3], Lateral Tripod [I4], Stick [I5] and Ventral [I6].

Precision category presents the following grasp types: Thumb-Index Finger [PC1], Inferior Pincer [PC2], Writing Tripod [PC3], Thumb-2 Finger [PC4], Thumb-3 Finger [PC5], Thumb-4 Finger [PC6], Tripod [PC7], Quadpod [PC8], Precision Sphere [PC9], Precision Disk [PC10], Tip Pinch [PC11] and Parallel Extension

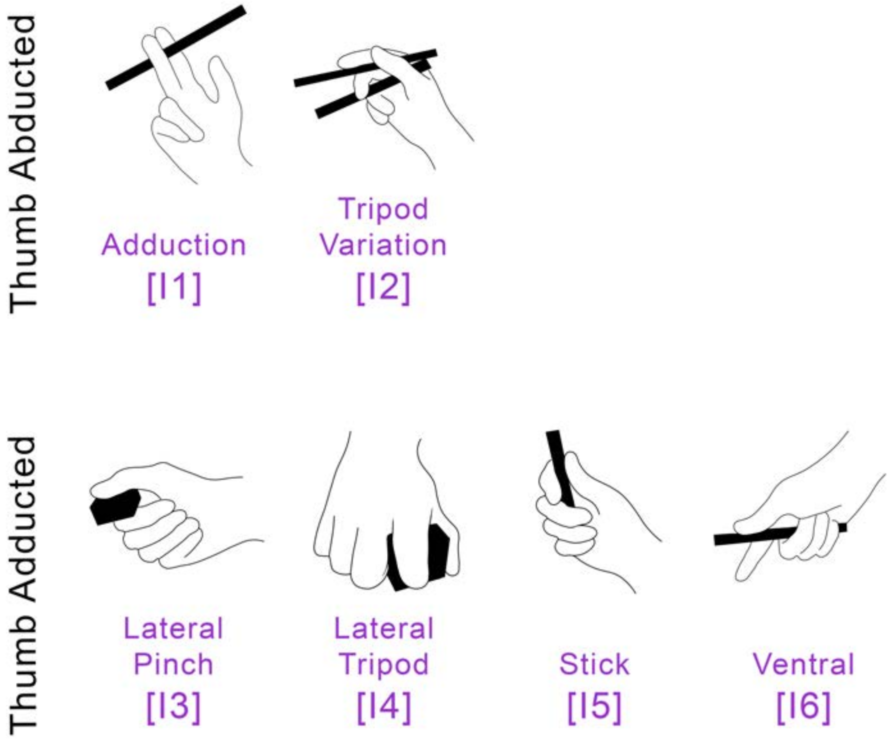


Figure 4.8: Intermediate Grasps from the Human GRASP Taxonomy (Feix et al., 2009)

[PC12].

4.6.3 Grasp Dimension

Feix et al. (Feix et al., 2014b) defined grasped dimension as the part of the object that lies between the fingers when grasped. By using the object axes (A, B and C), grasped dimension is analysed to indicate which axes best determine the hand opening. Figure 4.10 shows examples of how object dimensions determine the grasped dimension. In Figure 4.10 a) the object is grasped along the shortest

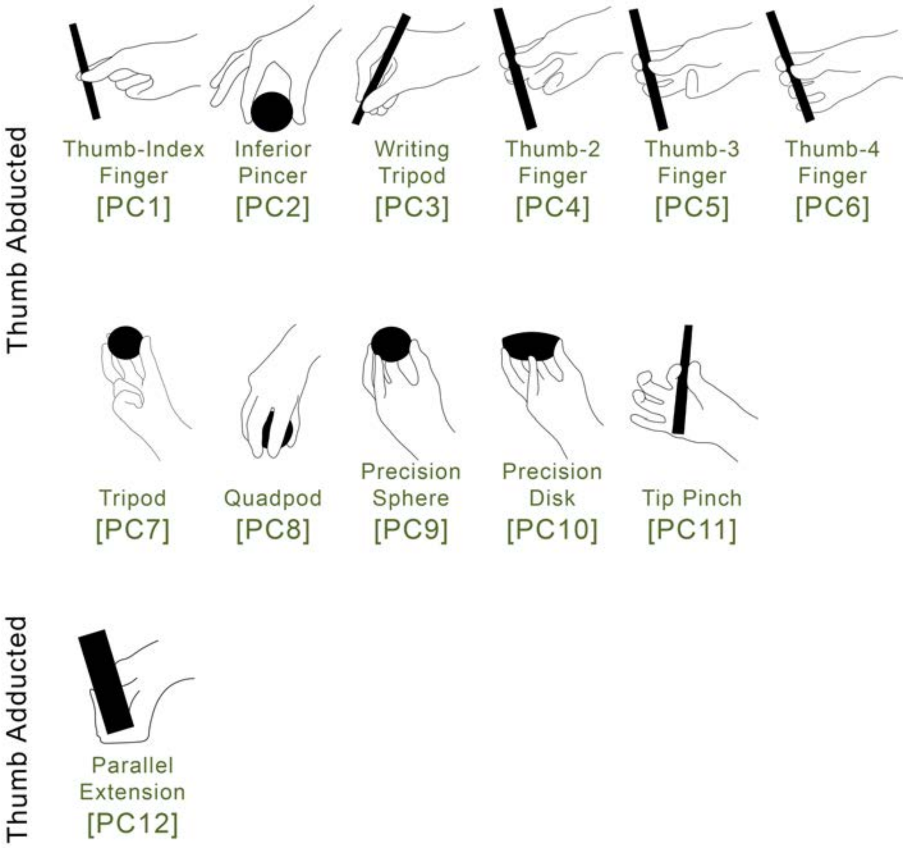


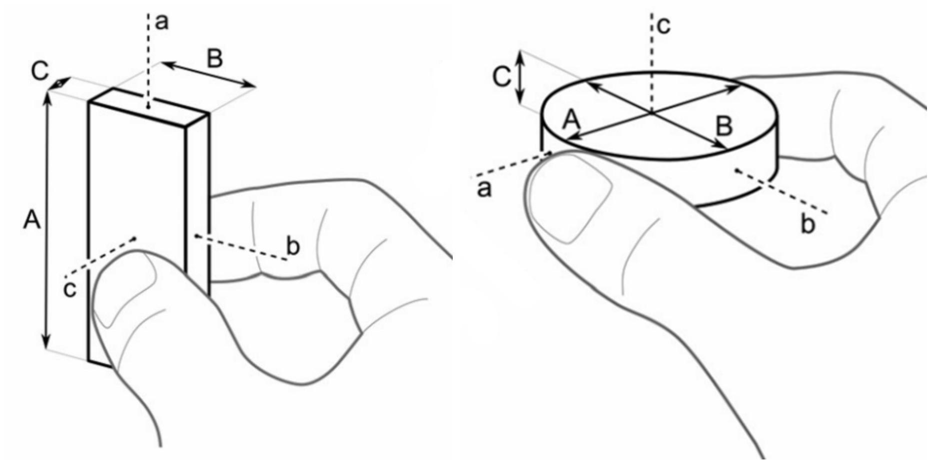
Figure 4.9: Precision Grasps from the Human GRASP Taxonomy (Feix et al., 2009)

dimension (dimension C) while in Figure 4.10 b) the object is grasped along dimension A/B, meaning that both A and B dimensions determine the hand opening.

4.7 Grasp Labelling

4.7.1 Methodology

In real grasping analysis, researchers annotated the grasps by analysing video recordings and making notes of the grasp characteristics (Feix et al., 2014b, 2014a;



(a) Grasp Dimension C (b) Grasp Dimension A/B

Figure 4.10: Grasped dimension examples as defined by Feix et al. (Feix et al., 2014b).

García Álvarez, Roby-Brami, Robertson, & Roche, 2017; Zheng, Rosa, & Dollar, 2011). Alvarez et al. (García Álvarez et al., 2017) conducted a study to identify and qualify grasp types used by patients with stroke and determine the clinical parameters that could explain the use of each grasp. For this, they video recorded patients with chronic stroke-related hemiparesis while grasping objects and asked two experienced observers to independently rate the type and quality of each grasp. When the results were different between the two observers, they re-analysed the video together until they reached a consensus. Similarly, Zheng et al. (Zheng et al., 2011) presented a study on the usage frequency of different grasp types throughout the daily functions of a professional housemaid and a machinist. They video recorded their hand usage during their daily work activities and further analysed the recording to report on grasp type as well as information related to the task and object involved in each grasp. One researcher went through the video recording and annotated the grasps using Cutokosky's grasp taxonomy (M. R. Cutkosky, 1989). This approach was also used for gesture elicitation analysis (Kipp, Neff, & Albrecht, 2007; Wobbrock et al., 2009; Martell, 2002), however this methodology can be time-consuming when large data set are analysed.

To address this, labelling in this work was done through a custom-made system to provide raters with an easy way to label the collected grasps, with data being automatically saved in CSV files ready for statistical analysis. For this, two academic raters labelled the grasps collected following the methodology of Feix et al. (Feix et al., 2014b, 2014a) for labelling real object grasps. The raters were academic members of staff with background in grasping literature and were trained to annotate the full set of grasps using grasp types, categories and dimensions using the common terminology from the Human GRASP Taxonomy (Feix et al., 2009) which is the current state of art in real grasp taxonomies. The custom-made system used for analysis is presented in the next section.

4.7.2 Custom-made Labelling System

To label the grasps collected during elicitation studies (images of the real view and the virtual view as described in Section 4.5.6), a custom-made system (labelling application) was developed. The application aimed to provide raters with a user interface for annotating the grasps collected during elicitation studies and was developed as a Windows Presentation Foundations (WPF) using C# programming language. WPF is a UI framework that creates desktop client applications and supports a broad set of application development features such as application model, resources, controls, graphics, layout, data binding, documents and security (Microsoft, 2021). Figure 4.11 shows the layout of the Labelling Application. When run, the application loads the virtual and real view for each grasp instance recorded during the elicitation study (as described in Section 4.5.6). The application shows a button for each grasp type from the Human GRASP Taxonomy (Feix et al., 2009) organised in the corresponding grasp classes (Power, Precision and

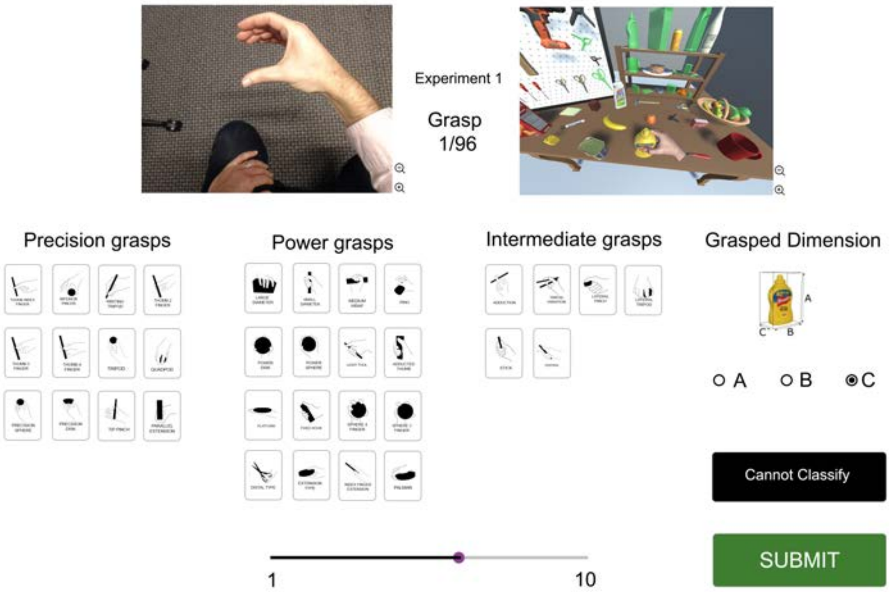


Figure 4.11: Labelling Application used for labelling grasp instances based on grasp category, type and dimension that was used for labelling grasp data collected in the experiments of this thesis.

Intermediate). The confidence level is presented as a slider from 1 to 10, which indicates the confidence in assigning a grasp type to a grasp instance. The application also shows a “Cannot Classify" button which is used if the virtual view and the real view do not provide a clear overview of the grasp performed (either due to occlusion or camera setup errors).

For cleaning and processing the grasp data and following current literature on grasp labelling methodologies (Feix et al., 2014b, 2014a), the following considerations were applied:

- If the confidence level was below five for at least one of the raters, that grasp instance was removed from the data set.
- If the grasp was labelled as “cannot classify" by at least one of the raters, the grasp was excluded from the data set.

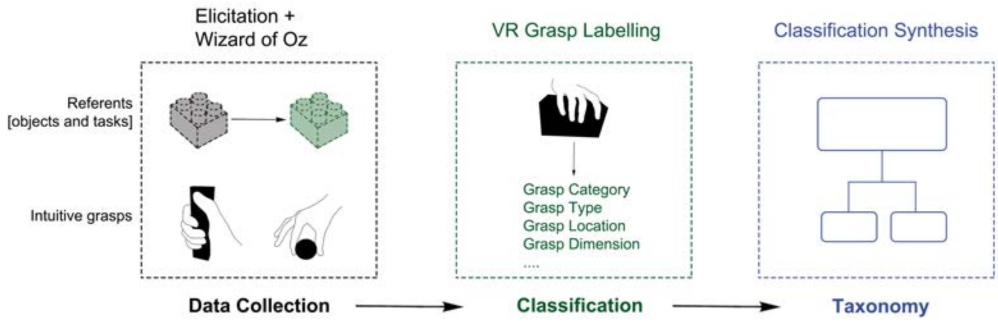


Figure 4.12: Overview of the method proposed for collecting grasps, classifying them based on current grasp metrics and synthesizing the results in the first VR Taxonomy of Grasp Types

Then, the “Submit” button is used when the rater is confident with the answers and ready to move to the next grasp instance for labelling. This is repeated until all grasp data has been annotated by both raters individually.

4.8 Ethical Approval

Ethical approval was obtained for all studies within this thesis following existing guidelines from Birmingham City University and prior to all studies taking place. All participants who took part in user experiments presented in this thesis signed a consent form and were informed about the details of the experiment. An example of the consent form can be found in Appendix A.

4.9 Conclusion

This chapter presented an overview of current methods for developing taxonomies in HCI, namely for data collection and classification. Inspired by these methods and methods for developing grasp taxonomies for real objects, a novel methodology is proposed. An elicitation combined with Wizard of Oz method is employed for collecting grasp data in the virtual environment, which is then classified based on current grasp metrics in a VR labelling application. These classifications are then synthesized in a VR taxonomy of grasp types. An overview of this method is shown in Figure 4.12.

The environment for collecting data is presented in detail in Section 4.5 which is used in virtual experiments presented in Chapters 5-8, with modifications in some components of the method based on the aim of each experiment. These modifications are presented in Table 4.3, where the method components are presented: Apparatus (Section 4.5.1), System architecture (Section 4.5.2), Virtual objects (Section 4.5.3), Tasks (Section 4.5.4), Grasp data (Section 4.5.6), Grasp classification (Section 4.6) and Grasp labelling (Section 4.7). Modifications are shown for each chapter and method component. "Standard" means that the method presented in this chapter was not modified.

Method Component	Chapter 5	Chapter 6	Chapter 7	Chapter 8
Apparatus	Standard (4.5.1)			
System Architecture	Standard (4.5.2)			
Virtual Objects	banana, mug, lego, marker, meat can, scissors, mustard	All objects in Sec. 4.5.3	Mug	All objects in Sec. 4.5.3
Tasks	Translate (1 axis)	Translate (3 axes)	Translate (1 axis)	Docking (6DOF)
Virtual Environment	Standard (4.5.5)			
Grasp Data	Standard (4.5.6)			
Grasp Classification	GAp Category Types	GAp Category Types	GAp Category Types Location	GAp Category Types Dimension
Grasp Labelling	Standard (4.7)			

Table 4.3: Overview of methodology used in the VR elicitation studies presented in this thesis. Standard refers to the fundamental method for each component of the methodology as presented in the subsections of this chapter.

5 | Virtual and Real Grasping

This work was published in the proceedings of 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) as "A Grasp on Reality: Understanding Grasping Pattern for Object Interaction in Real and Virtual Environments" (Blaga et al., 2021b).

5.1 Introduction

Freehand grasping is a current challenge in VR, with grasp models taking knowledge from real grasping research to develop intuitive interactions based on the assumption that humans interact with virtual objects in a similar manner to how they interact with objects in reality. However, researchers showed that there are differences in spatial perception (B. Williams, Narasimham, Westerman, Rieser, & Bodenheimer, 2007) as well as hand movements for the same tasks (Viau, Feldman, McFadyen, & Levin, 2005) between Real Environments (RE) and Virtual Environments (VE). These differences led to a more detailed investigation on knowledge transfer in interaction between REs and VEs, with researchers showing that the lack of haptic feedback introduces a switch from real-time visual control to one that depends on visual perception, which in VR differs based on hardware limitations (Murcia-López & Steed, 2018). In spite of these limitations, VR interactions that mimic real life interactions have shown to influence users to partially transfer knowledge from real to virtual environments (Multon & Olivier, 2013), therefore it can be assumed that users would try to mimic real grasping interactions in VR, with some differences being introduced due to the existing limitations. However, a comparison between freehand grasping movements in REs and VEs has not been conducted yet. To address this, this chapter presents the first analysis of differences in grasping patterns when interacting with real and virtual objects. A user elicitation study was conducted where participants (N = 20) were asked to grasp and translate seven physical objects. This experiment was then replicated in VR, using 3D representations having the same size, shape and texture as the physical objects used in the real environment. Grasps collected were analysed based on grasp aperture and grasp labels with results presenting an overview of similarities and differences between grasping interaction metrics for real and virtual environments.

The chapter is structured as follows: Section 5.2 presents a literature review of existing work on comparisons between RE and VE; Section 5.3 presents the experiment design with a detailed methodology being presented for selecting apparatus, objects and tasks; Section 5.4 presents the protocol for the user elicitation study presented in more detail in Chapter 4; Section 5.5 presents the metrics used for analysing grasping patterns; Section 5.6 presents the proposed hypothesis; Section 5.7 presents the methodology used for data analysis; Section 5.8 presents results where grasp aperture (GAp) and grasp labels are analysed and compared for real and virtual objects and Section 5.9 presents discussion and conclusions which summarizes the similarities and differences in grasping patterns identified between real and virtual objects.

5.2 Background

An important application of virtual environments is based on the assumption that what people learn from exploring physical environments is functionally similar to what they learn from exploring virtual renderings of them. However, this might not always be true due to the nature of VR, which led researchers to focus on comparing virtual and real environments for improving existing systems and understanding VR parameters such as spatial presence (B. Williams et al., 2007), training transfer (Murcia-López & Steed, 2018) and hand movements (Levin et al., 2008). Understanding these differences has been of high importance for validating virtual models and understanding how VR technology can be improved for natural and intuitive use (Murcia-López & Steed, 2018). For example, Plumert et al. (Plumert, Kearney, Cremer, & Recker, 2005) investigated differences between distance perception in real and virtual environments. To achieve this, they conducted user experiments where participants estimated how long it would take them to walk to targets in real and virtual environments. Their findings show that people underestimated time to walk in both environments, with no statistically significant differences being observed between real and virtual environments. However, another study performed by Ziemer et al. (Ziemer, Plumert, Cremer, & Kearney, 2006) showed that when participants are asked to estimate distances in the VE first, estimates were significantly shorter in VE than in RE. Other works also analysed distance perception in VE and RE and found significant differences between the environments, mainly caused by the small FOV (B. Wu, Ooi, & He, 2004) and weight (Willemsen, Colton, Creem-Regehr, & Thompson, 2004) of the HMDs. Spatial perception has also been investigated by Williams et al. (B. Williams et al., 2007) who showed that participants took significantly longer to accomplish a change in perspective in the VE than in the RE, which might have been due to poorer quality in rendering the VE in the HMD (as opposed to vision of the RE) or the limited resolution of the HMD which affected the sense of immersion and therefore led to weaker spatial representations in VE.

Since distance perception and spatial knowledge play an important role in the day-to-day interactions in real environments, these differences shown in VEs might also influence the way users interact with virtual objects, which led researchers to investigate hand movements in VE and RE to allow development of improved virtual rehabilitation systems (Viau et al., 2005). Viau et al. (Viau et al., 2005) showed that hand movements in VE are similar to hand movements in RE, however, some differences were shown in wrist and elbow movements which might be linked to the lack of feedback at the end of the grasp. To address this, Levin et al. (Levin et al., 2008) analysed the effect of haptic feedback on movement patterns in VEs and addressed the need for identifying movements made in VEs and how similar or different these movements are for allowing improved development of training systems. For this, they asked participants to perform three grasps in RE and VE using a data glove and found that movements were slower and had longer deceleration times, elbow extension was greater and apertures were wider in the VE. The authors suggest that differences in movement kinematics were likely due to the lack of prior experience with the VE, an uncertainty of object location (which is directly linked to spatial perception (B. Williams et al., 2007)) and limited FOV in the HMD as in the work of (B. Wu et al., 2004). Moreover, the authors show that movements were significantly slower when users wore the data glove as opposed to freehand interaction in RE. Limitations with glove-based systems for natural interactions have been identified before and are presented in more detail in Chapter 2. To address this, Whitwell et al. (Whitwell, Ganel, Byrne, & Goodale, 2015) focused on understanding how removing augmentation of the hand and haptic feedback from a natural grasping task influences the interaction approach and found that the lack of haptic feedback introduces a switch from real-time visual control towards one that depends more on visual perception and cognitive supervision, which has shown to be different in VE (Murcia-López & Steed, 2018). Considering these differences in perception between REs and VEs, it can be assumed that differences in grasping patterns such as grasp type and grasp category between RE and VE exist, which might be the root of the challenges with current grasp models presented in Chapter 2. However, a direct comparison between RE and VE grasping patterns has not been conducted yet and therefore these differences are still unknown. This study aims to address this gap and explore a comparison between RE and VE grasping patterns.

5.3 Experiment Design

To analyse and compare grasping patterns in VE and RE, a user study (N = 20) was conducted on seven representative real objects and their virtual twins from the *"Yale-Carnegie Mellon University-Berkeley Object and Model Set"* (Calli et al., 2015). This section provides a detailed overview of the virtual objects, task, apparatus, environment and 3D hand model used for interaction as detailed in Table 4.3.

5.3.1 Objects

Chapter 4 describes the methodology for selecting 16 virtual objects from the *Yale-Carnegie Mellon University-Berkeley Object and Model Set* (Calli et al., 2015) which are used in the experiments of this thesis. From this set, taking into consideration the need for replicating these objects in the real environment for a direct comparison of grasping patterns in RE and VE, a subset of seven objects is selected: Banana, Mug, Lego, Marker, Meat Can, Scissors and Mustard (see Figure 5.1). When selecting this subset, the aim was to select objects that fit in multiple daily object categories (Calli et al., 2015) as detailed in Chapter 4: Food items, Kitchen items and Tool items. While one important limitation in selecting this dataset was replicating them in the real environment, a notable factor in choosing this subset was ensuring diversification of objects that participants are familiar with while including a variety of shapes (Banana - cylindrical, Lego - cube, Meat Can - rounded cuboid, Mustard - irregular shape), sizes (from 26 mm to 200 mm) and objects that propose a range of manipulation uses (Mug, Marker, Scissors).

These selected objects are presented in Figure 5.1 along with the object dimensions in mm. To ensure size is consistent across RE and VE conditions, the 3D models are resized to match the size of the physical objects, using scale-related settings in Unity (2018.2).

(a) Banana (b) Mug

(c) Lego (d) Marker

(e) Meat Can (f) Scissors

(g) Mustard

Figure 5.1: Objects chosen for the study with dimensions. The objects were chosen from the *Yale-Carnegie Mellon University-Berkeley Object and Model Set*, which present the most frequently used objects in research (Calli et al., 2015).

5.3.2 Task

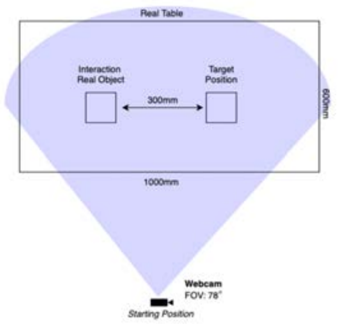
To evaluate the grasping patterns in RE and VE, an object translation task on the *X* axis of the Cartesian coordinate system was implemented in the positive direction for both experimental conditions (RE and VE as shown in Figure 5.2). The task was consistent across conditions, with participants being asked to move the object to the target position, which was positioned 300 mm away from the object to be grasped in both RE and VE (see Figure 5.2).

5.3.3 Apparatus

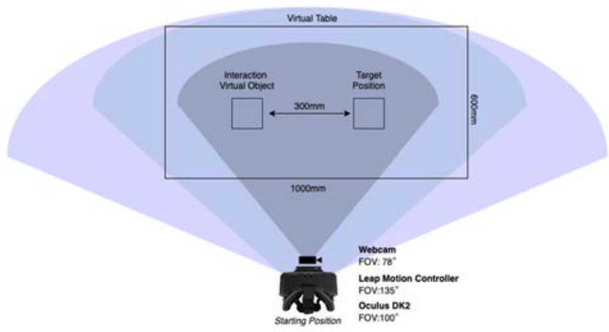
The grasps performed during the task were recorded in both environments using a Logitech Pro 1080p HD camera with a Field of View (FOV) of 78°as shown in Figure 5.2. Pilot tests were conducted to find the optimal position for attaching the camera to facilitate recording of participants' hands at all times. The following configurations were used for RE and VE conditions:

RE: Participants wore a camera attached on their forehead using a head strap (GoPro Head Strap), following the methodology used in real object grasping research (Bullock et al., 2013; Feix et al., 2014b, 2014a) where subjects wore a head-mounted camera to record their grasping movements during experiments. The camera was centred and tilted by 30° to record participants' hands during the interaction. The starting position was the same for all participants.

VE: Participants wore the Oculus DK2 VR headset, the Leap Motion device and the camera as detailed in Chapter 4. The Leap Motion Controller was attached to the HMD, facing the user's hands. The camera was then attached on top of the Oculus DK2 facing participants' hands and recording all the grasps during the VE experiment. The virtual interaction space was 600 (mm) $\times$ 600 (mm) $\times$ 600



(a) RE Experimental Environment



(b) VE Experimental Environment

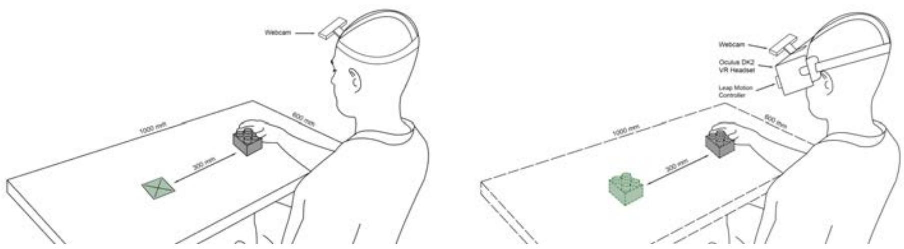
Figure 5.2: Experimental Environment; a) RE Experimental Environment consisted of the Logitech Webcam, with a FOV of 78°. The physical table was 600 mm × 1000 mm, with the physical objects positioned on it, 300 mm away from the target position. The starting position was consistent for both (a) VE and (b) RE Experimental Environments. b) VE Experimental Environment consisted of the Oculus DK2, with the Leap Motion Controller and Logitech Webcam attached to the HMD. The virtual table was 600 mm × 1000 mm, with the virtual objects positioned on it, 300 mm away from the target position. The webcam had a FOV of 78°, the Leap Motion Controller a FOV of 13°, and Oculus DK2 a FOV of 100°.

(mm) (based on Leap Motion Controller FOV). The system was developed using C#, Unity 2018.2 and the Leap Motion 4.0 SDK. The starting position was the same for all participants.

5.3.4 Environment

The user experiment was conducted in a controlled environment under laboratory conditions. The test room was lit by a 2700k (*warm white*) fluorescent with no external light source.

The RE was composed of a real table of 600 mm × 1000 mm and the real objects as shown in Figure 5.2 a. The location where the real objects had to be translated



(a) Real Environment (b) Virtual Environment

Figure 5.3: Experiment environment for the two conditions: a) RE shows the participant wearing the head-mounted camera, seated in front of the physical table and grasping a real object. The green marker represents the position where the participant needs to move the real object. b) VE shows the participant wearing the VR equipment, seated in front of a virtual table. The green virtual marker represents the position where the participant needs to move the virtual object.

to was indicated by a marker on the table, as shown in Figure 5.3.

The VE was composed of a virtual table of 600 mm $\times$ 1000 mm and the virtual objects as shown in Figure 5.2 b). The location where the virtual objects had to be translated was indicated by a marker represented by a replica of the virtual object coloured in green as shown in Figure 5.3 and detailed in Chapter 4.

5.3.5 Hand Representation

The hand model used in this experiment is an abstract hand model from the Leap Motion SDK, represented as a set of cylinders and spheres representing bones and joints. The abstract hand model was chosen due to being easily accessible and avoiding noticeable gender characteristics (Schwind et al., 2017) as well as being prominently used for interaction studies (Cohen, Voldman, Regazzoni, & Vitali, 2018; Tayag, Ronie, & Marvin, 2021).

5.3.6 Participants

A total of 20 right-handed participants (12 males and 8 females), ranging in age from 19 to 65 (M = 33.25, SD = 11.98) and from a population of university students and staff members volunteered to take part in this study. Participants were asked to self-assess their level of experience with VR systems, with 6 participants reporting to have an average level of experience, 11 reported being novice to the technology and 3 self-labelled themselves as experts. Participants did not have any previous experience with hand tracking sensors. All participants completed both conditions of the experiment and a standardised consent form. Visual acuity of participants was measured using a Snellen chart. Each participant was also required to pass an Ishihara test to check for colour blindness. Participants with colour blindness and/or non corrected visual acuity of < 0.80 (where 20/20 is 1.0) were not included in this study. Participants were not compensated.

5.4 Protocol

5.4.1 Training

Participants underwent initial hand interaction and task training to familiarise themselves with the environments. The training task was a representative version of the tasks in the user study, where participants were asked to grasp and translate a cube object, both in RE and VE. For this, a physical cube was used in RE and a 3D virtual representation of the cube in VE (50 mm $\times$ 50 mm $\times$ 50 mm).

5.4.2 Test

Each participant performed both conditions: RE and VE. Half participants started with RE and the other half with VE. The 7 objects were randomised for each condition with the starting position being consistent for all objects. Each participant grasped every object three times, with a total of 21 grasps (7 objects x 3 repetitions) performed per participant, with a total of 840 grasps collected in RE and VE. Navigation was not considered in this work, therefore users were seated during the experiments as in other user elicitation studies (Piumsomboon et al., 2017) and the work of (Bozzacchi, Volcic, & Domini, 2014) exploring grasping movements. Moreover, it has been shown that sitting in VR naturally induces a strong sense of orientation in the virtual environment (Peillard et al., 2019). Participants were allowed to move to adjust their perspective if needed.

RE: Participants were wearing the head mounted camera as presented in Section 5.3.3. Chapter 4 details the use of Wizard of Oz methodology for the VR experiments presented in this thesis. While a Wizard of Oz can not be fully replicated in RE, to ensure consistency in data collection between RE and VE, a Wizard of Oz-like study was implemented in RE, where users were asked to indicate when they were happy with their grasp, before completing the translation task.

For each task, participants were instructed to be seated in a neutral position with their hands at the side, while the test coordinator placed the object under experiment in the starting position. Objects were placed in front of participants in randomised order. The test coordinator then informed participants when they could start the task. Participants then started the grasping task and informed the test coordinator when they were comfortable with their grasp. At this point the hand pose was recorded using the web camera. When participants finished the translation task, they would return in neutral position while the test coordinator prepared the next object under experiment.

VE: Participants were wearing the Oculus DK2 as presented in Section 5.3.3. As detailed in Chapter 4, the Wizard of Oz methodology was used. At the beginning of each task, a virtual object appeared on the virtual table. For each task, participants were instructed to be seated in a neutral position with their hands at the side, while the virtual object under experiment appeared in the starting position. Virtual objects appeared in randomised order. The test coordinator instructed participants to grasp the virtual object in the way it felt most intuitive, notifying the test coordinator when they were comfortable with the grasping position, which led the test coordinator to trigger the interaction between the virtual hand and object and therefore allowing participants to move the object to the required position. When participants finished the translation task, they would return in neutral position until another virtual object appeared in the starting position.

5.5 Metrics

Grasp Aperture (GAp): GAp is a metric used to measure how accurately users estimate the size of a virtual object and has been described in more detail in Chapter 4.

Grasp Labels: Grasp Labels refers to labels assigned to grasp instances (grasp category and grasp type) during the Labelling process as detailed in Chapter 4. The grasp categories are Power, Intermediate and Precision. Power grasps are linked to stability and security. These grasps are distinguished by large areas of contact between the hand and the object (M. R. Cutkosky, 1989) as shown in Figure 5.4. Intermediate grasps present elements of *Power* and *Precision* roughly

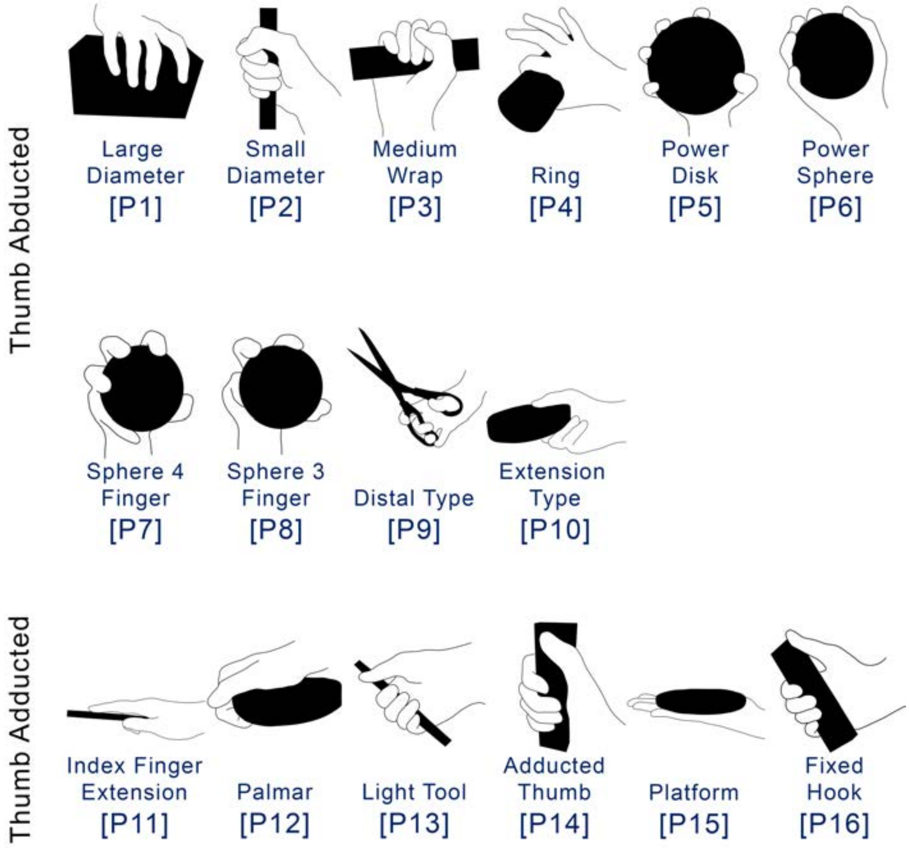


Figure 5.4: Power Grasps

in the same proportion, enabling a finer representation of grasp types (Bullock et al., 2013) as shown in Figure 5.5. Precision grasps present grasps where the object is commonly held between the fingertips. While this allows an increased level of manipulation by movement of the fingertips, the object cannot be gripped firmly (Lyons, 1985) as shown in Figure 5.6. Figures 5.4, 5.5, 5.6 present grasp types for each grasp categories, showing the name of each grasp and an ID in brackets, which are presented in more detail in Chapter 4 and are used for reporting grasp type results.

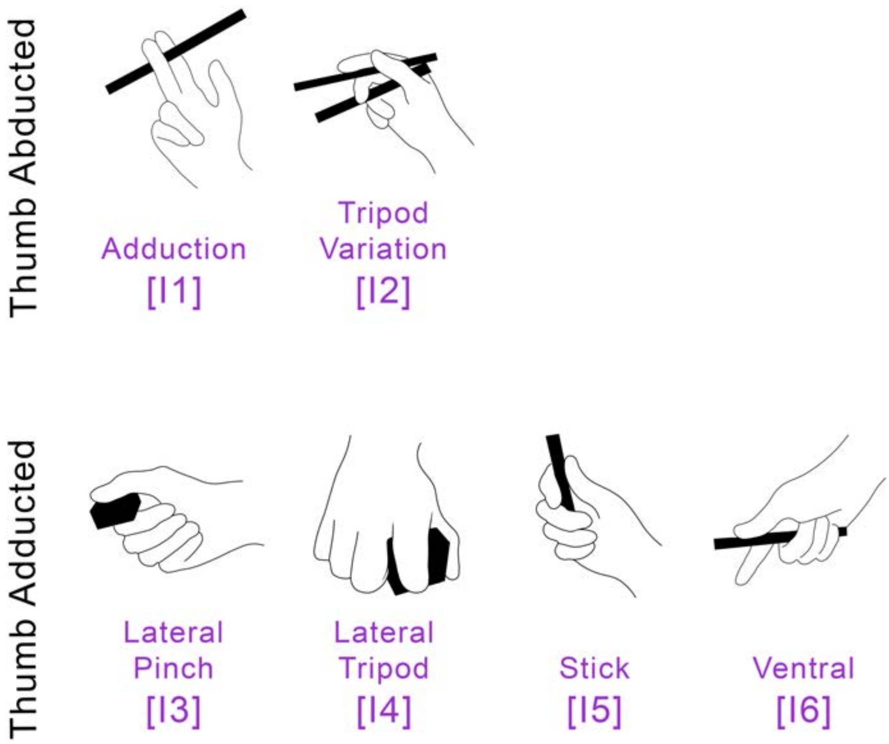


Figure 5.5: Intermediate Grasps

5.6 Hypothesis

Following the current literature defined in this chapter that suggests differences in visual perception and hand movement between real and virtual interactions, the following hypothesis is proposed: **Grasping patterns for interacting with virtual objects are different than interacting with real objects.**

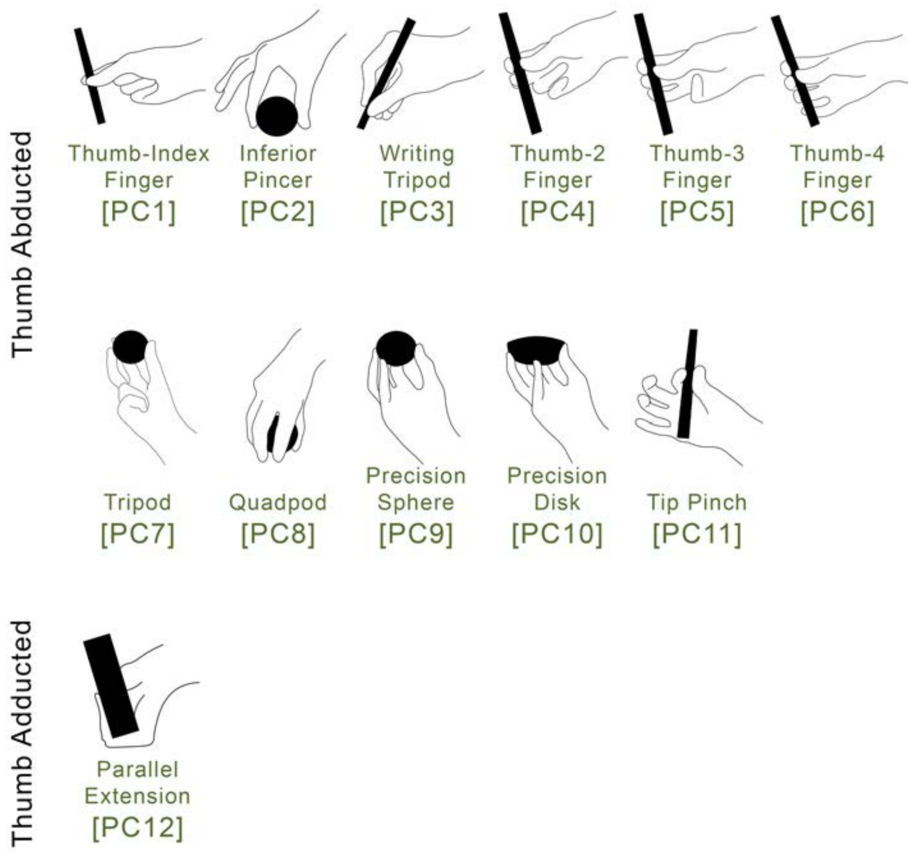


Figure 5.6: Precision Grasps

5.7 Data Analysis

The Shapiro-Wilk (Shapiro & Wilk, 1965) normality test found the data to be not normally distributed. For testing statistical significance between grasp patterns in RE and grasp patterns in VE, where the dependent variable (Grasp Category) is nominal categorical (Power, Intermediate and Precision), contingency tables were created and analysed for significance using a Chi-Squared Test of Independence with 95% Confidence Intervals, therefore, a p-value of less than .05 will indicate statistical significance. Cramer's V calculation for effect sizes was applied after verifying its assumptions that the variables under analysis are categorical. Results were interpreted based on existing guidelines based on degrees of freedom.

5.8 Results

5.8.1 Grasp Aperture (GAp)

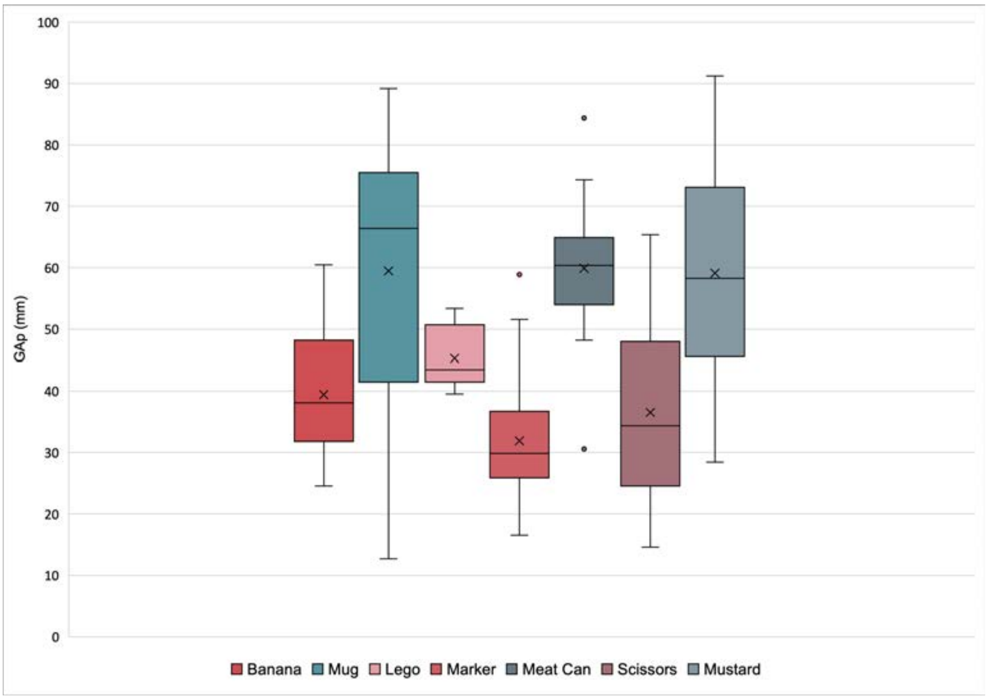


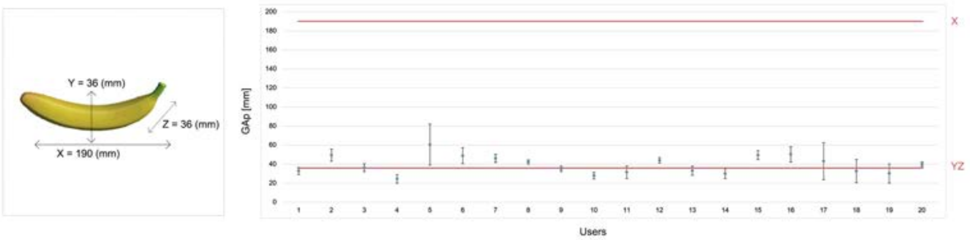
Figure 5.7: GAp in mm for virtual objects used in this experiment. X marks on boxplots indicate the mean GAp across all participants for objects used in this experiment (Banana, Mug, Lego, Marker, Meat Can, Scissors, Mustard). The line that divides the box in two plots shows the median, which marks the mid-point of the data. Whiskers represent the highest and lowest values within 1.5 times the interquartile range. Outliers are shown in coloured circles.

Grasp Aperture (GAp) is defined as the distance between the thumb and the index finger and is reported for understanding how accurately users estimate the size of the virtual objects. Figure 5.7 shows GAp for every virtual object used in this experiment, which ranged from 31.89 mm to 59.89 mm. Mean GAp and standard deviation are reported for individual objects as follows: Banana (M = 39.41, SD = 9.42), Mug (M = 59.48, SD = 22.41), Lego (M = 45.25, SD = 4.80), Marker (M = 31.89, SD = 10.09), Meat Can (M = 59.89, SD = 11.10), Scissors (M = 36.53, SD = 13.71), Mustard (M = 59.12, SD= 17.70). Figure 5.8 shows for every virtual object the mean GAp in mm of every participant (N = 20) with Standard Error (SE). Object size on X,Y,Z are plotted as a reference for understanding how accurately participants estimated object size.

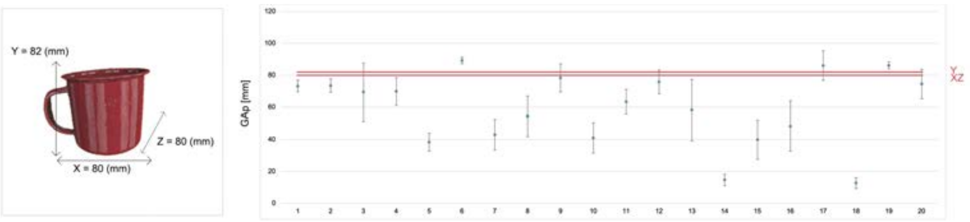
5.8.1.1 Analysis - GAp

For an accurate comparison in terms of GAp between real and virtual grasping, GAp needs to be measured in both environments. Since this experiment aimed to replicate the Wizard of Oz methodology in both environments without augmenting the hand for interaction, GAp was not measured in the real environment. However, it is known that in real grasping GAp is influenced by object size (Feix et al., 2014b), humans adjusting GAp based on object measurements to perform a stable grasp. Moreover, researchers have shown that in real environments humans grasp objects across their smallest dimension.

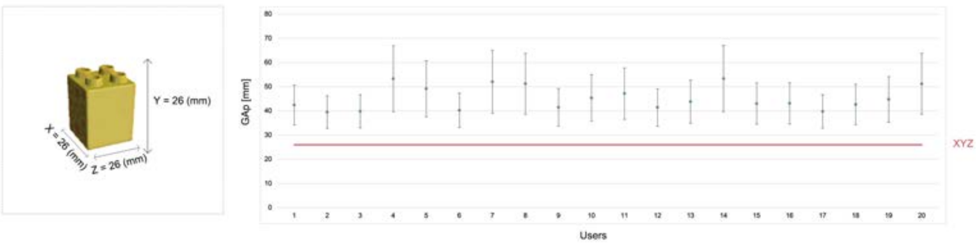
Considering this, it can be assumed that in a real environment, a banana object, which has the smallest dimension equal to 36 mm, would be grasped with a GAp of 36 mm to create a stable grasp. However, results of this experiment showed high variability between participants in choosing a GAp for a virtual banana, as shown in Figure 5.8 a), with participants grasping it larger and smaller than the dimension of the object. An example of these grasps is shown in Figure 5.9. In Figure 5.8 a), it is shown that User 5 performed a GAp larger than object size on XY, which is exemplified in Figure 5.9 a, User 20 performed a GAp approximately equal to object size, which is exemplified in Figure 5.9 b and User 14 performed a GAp smaller than object size, also known as interpenetration between the hand



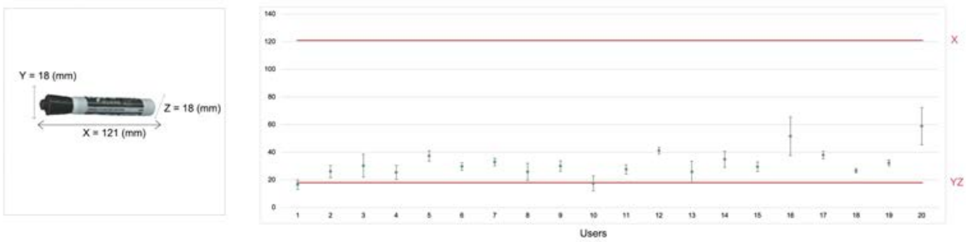
(a) Banana GAp



(b) Mug GAp

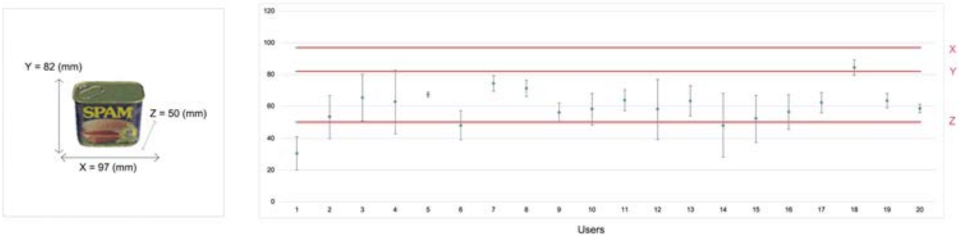


(c) Lego GAp

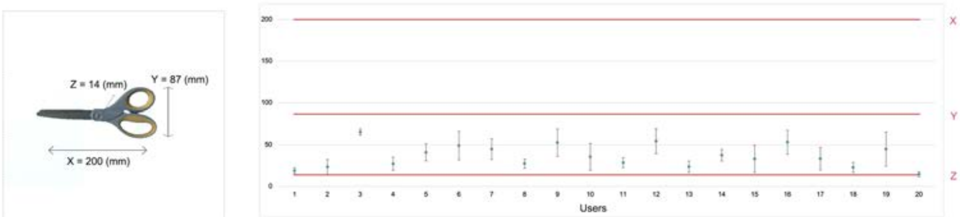


(d) Marker GAp

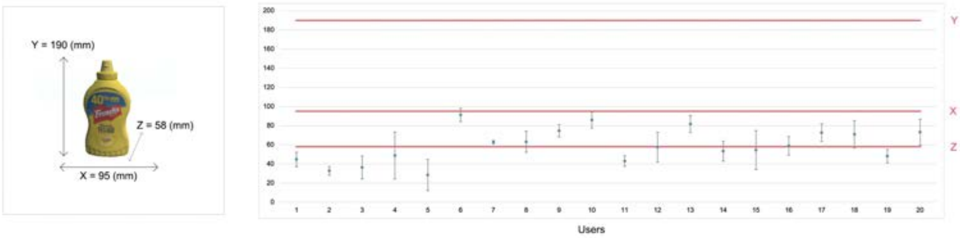
Figure 5.8: GAp for individual objects. Object dimensions are presented for each object in mm and plotted as red lines in the point graph. Green points represent the mean GAp for each participant in the user experiment (N = 20) with Standard Error (SE) bars.



(e) Meat Can GAp



(f) Scissors GAp

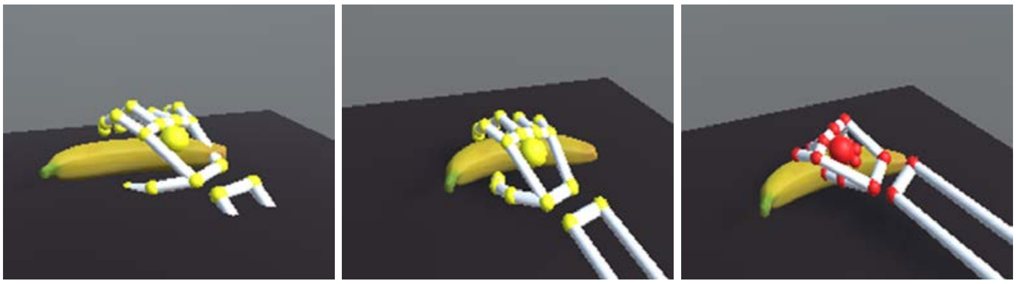


(g) Mustard GAp

Figure 5.8: GAp for individual objects. Object dimensions are presented for each object in mm and plotted as red lines in the point graph. Green points represent the mean GAp for each participant in the user experiment (N = 20) with Standard Error (SE) bars.

and the object in VR literature (Borst & Indugula, 2005), which is exemplified in Figure 5.9 c.

This high variability in GAp was observed for all objects used in this experiment, with unique patterns being identified for specific objects, which is contrary to Mixed Reality (MR) literature which showed that for grasping simple virtual objects (cubes, spheres) users chose a common GAp, irrespective of object size or shape (Al-Kalbani et al., 2016a). This might be due to users mimicking grasping



(a) User 5 (b) User 20 (c) User 14

Figure 5.9: Grasp examples from users to show a GAp a) larger, b) approximately equal and c) smaller than object size. User IDs with the GAp can be found in GAp graphs shown in Figure 5.8 on the X axis.

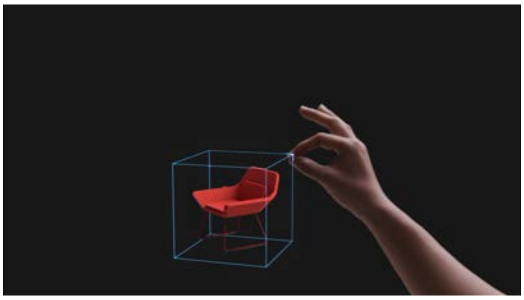


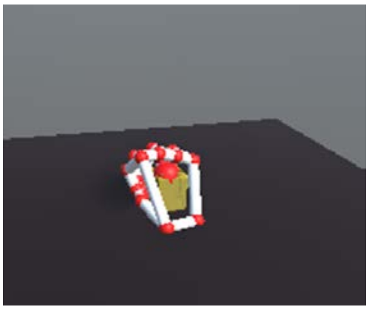
Figure 5.10: Pinch Grasp example from Microsoft Hololens 2 Docs

behaviour from real environments in VR and attempting to match the size of the object for daily life objects. However, the lack of haptic feedback at the end of the grasp might have introduced errors in object size estimation, which might be the cause of users grasping objects larger and smaller than their size. This is consistent with interaction literature showing that when removing haptic feedback from a natural grasping task, grasping decision is made solely based on visual perception and cognitive supervision, which introduces errors in size estimations (Whitwell et al., 2015; Murcia-López & Steed, 2018).

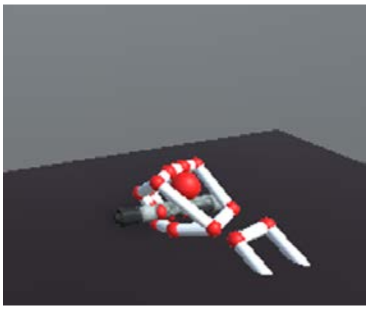
Objects of a more complex shape, for example the Mug, which can be grasped using the handle, the body or the top of the object showed higher variability in GAp with participants grasping smaller and larger, possibly due to users grasping the object on different locations and adjusting their GAp to match the size of the graspable body. Smaller objects such as Lego and Marker showed a different pattern, being grasped predominantly larger than object size. In RE, when the hand occludes the object, the haptic feedback at the end of the grasp guides the estimation of object size. However, in VR, when the hand occludes the object (see Figure 5.11) and there is no haptic feedback at the end of the grasp, the users' ability to estimate object size with accuracy is compromised. This error in size estimation could lead to larger and smaller grasps, however in this work, smaller objects were grasped larger in all collected instances (see Figure 5.8 c). This might be due to users avoiding performing a common "pinch grasp" associated with virtual object interaction (Balani & Tümler, 2021) where the thumb and the index finger come together to represent a virtual grasp (Figure 5.10) and instead trying to mimic real grasping interaction where fingers are wrapped around the object to perform a stable grasp. Moreover, MR research has shown that the most comfortable virtual object grasp size is equal to 80 mm regardless of object size (Al-Kalbani et al., 2016a), which might explain the larger apertures for smaller objects, users aiming to perform a comfortable grasp instead of accurately matching the size of the virtual object. However, to fully understand grasping patterns in VR and their differences to real grasping, the next section explores the positions of the fingers and palm, also known as grasp labels.

5.8.2 Grasp Labels

A total of 840 grasps (2 environments x 20 participants x 7 objects x 3 repetitions) were labelled during this experiment, following the methodology presented in Chapter 4. Out of 840 grasps, 14 grasps (4 grasps for RE and 10 grasps for VE) were removed due to being rated as "Cannot Classify" by at least one of the raters. The remaining 826 were analysed for understanding grasping patterns in RE and



(a) Hand occlusion when grasping Lego



(b) Hand occlusion when grasping Marker

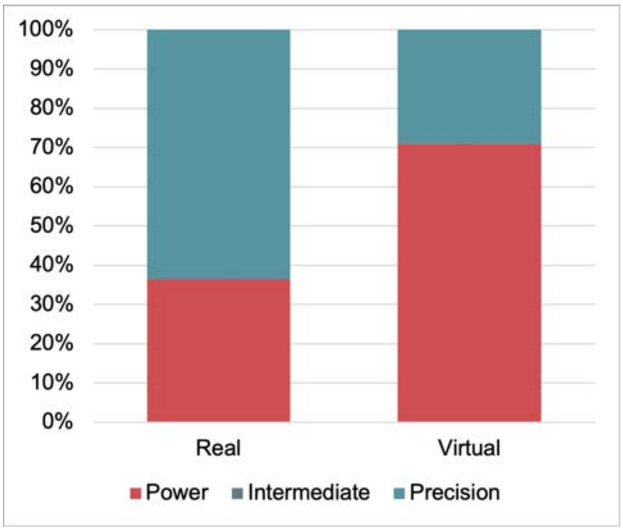
Figure 5.11: Hand occlusion examples in VR grasping.

VE. Cohen's Kappa was used to measure inter-rater reliability for labelling the grasps for both RE and VE. Raters agreed in 83% of instances (Cohen's Kappa = 0.41) which based on existing guidelines is a moderate agreement, often achieved when subjectivity is involved in the process (S. Sun, 2011) and is a common agreement score for classification tasks (Feix, Bullock, & Dollar, 2014c).

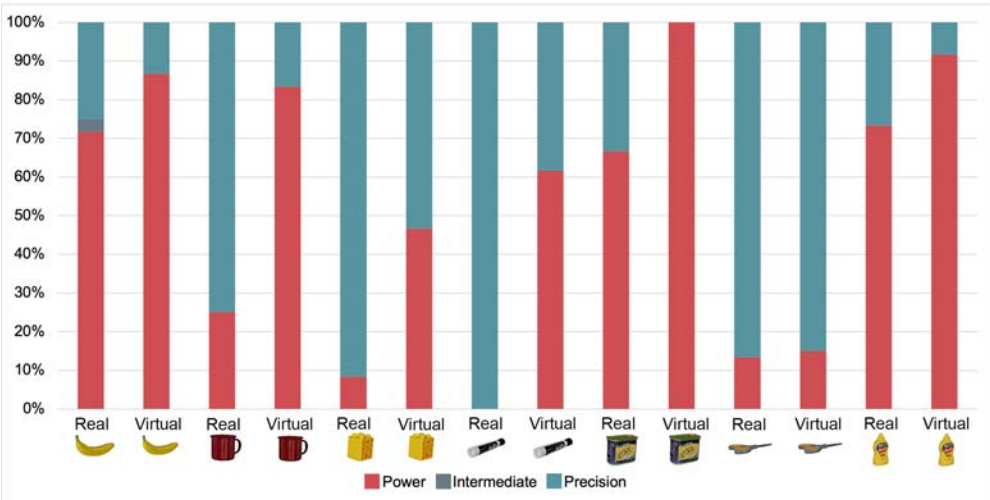
5.8.2.1 Grasp Category

Figure 5.12 a provides an overview of grasp categories (Power, Intermediate and Precision) used in RE and VE. In RE users grasped objects using Precision grasps in 63.22% (N = 263) instances, Power grasps in 36.29% (N = 151) instances and Intermediate grasps in 0.48% (N = 2) instances. In VE users grasped objects using Power grasps in 70.97% (N = 291) instances, Precision in 29.02% (N = 119) instances and Intermediate grasps in 0% instances. A Chi-Squared Test of Independence showed that this difference was statistically significant χ^2 **(2, N = 826) = 95.01, p < .001*** with medium ES (Cramer's V = 0.336).

Statistical significance was also tested for individual objects in terms of grasp category labels (RE compared to VE) with 95% Confidence Intervals. Figure 5.12 b) provides an overview of grasp categories (Power, Intermediate and Precision)



(a) Overall RE and VE Grasp Categories



(b) RE and VE Grasp Categories for individual objects

Figure 5.12: Power, Intermediate and Precision grasps ratio shown for overall RE and VE and for individual objects in RE and VE.

used in RE and VE for individual objects. Statistical significance was found for Mug RE (Precision 76.26% N = 45; Power 23.72% N = 14; Intermediate: 0%, N = 0) compared to Mug VE (Power 86.20%, N = 50; Precision 13.79% N = 8; Intermediate 0%, N = 0): χ^2 **(2, N = 117) = 35.53, p < .001*** with a large ES (Cramer's V = 0.54).

Lego RE (Precision 91.66%, N = 55; Power 8.33%, N = 5; Intermediate 0%) compared to Lego VE (Precision 53.33%, N = 32; Power 46.66%, N = 28; Intermediate 0%): χ^2 **(2, N = 120) = 31.34, p < .001*** with a large ES (Cramer's V = 0.51).

Marker RE (Precision 100%, N=57; Power 0%, Intermediate 0%) compared to Marker VE (Power 61.66%, N = 37; Precision 38.33%, N = 23; Intermediate 0%): χ^2 **(2, N = 117) = 48.23, p < .001*** with a large ES (Cramer's V = 0.63).

Meat Can RE (Power 66.10%, N = 39; Precision 33.89%, N = 20; Intermediate 0%) compared to Meat Can VE (Power 100%, N = 60; Precision 0%; Intermediate 0%): χ^2 **(2, N = 119) = 20.91, p < .001*** with a large ES (Cramer's V = 0.41).

Mustard RE (Power 74.57%, N = 44; Precision 25.42%, N = 15; Intermediate 0%) compared to Mustard VE (Power 93.22%, N = 55; Precision 6.77%, N = 4; Intermediate 0%): χ^2 **(2, N = 120) = 7.01, p = .031*** with medium ES (Cramer's V = 0.24).

No statistical significance was found for Banana RE (Power 71%, N = 42; Precision 25.42%, N = 15; Intermediate 3.38%, N = 2) compared to Banana VE (Power 89.65%, N=52; Precision 10.34%, N = 6; Intermediate 0%): χ^2 *(2, N = 117) = 4.18, p = .123*. Scissors in RE (Precision 86.66%, N = 52; Power 13.33%, N = 8; Intermediate 0%) compared to Scissors in VE (Precision 86.20%, N = 50; Power 13.79%, N = 8; Intermediate 0%):χ^2 *(2, N = 118) = 3.34, p = .188*.

Object	RE			VE			Statistic
	P	PC	I	P	PC	I	
Mug	23%	76%	0	86%	13%	0	**Stat = 35.53, p < .001***
Lego	8%	91%	0	46%	53%	0	**Stat = 31.34, p < .001***
Marker	0	100	0	61%	38%	0	**Stat = 48.23, p < .001***
Meat Can	66%	33%	0	100%	0	0	**Stat = 20.91, p < .001***
Mustard	74%	25%	0	93%	6%	0	**Stat = 7.01, p = .031***
Banana	71%	25%	3%	89%	10%	0	*Stat= 4.18, p = .123*
Scissors	13%	86%	0	13%	86%	0	*Stat = 3.34, p = .188*

Table 5.1: Grasp category results for objects presented in this study, in both RE and VE conditions. Percentages for use of each grasp category (P for Power, PC for Precision and I for Intermediate) are shown for each object together with the statistical results for comparing between conditions.

5.8.2.2 Most Common Grasp Types

Table 8.6 shows an overview of the most used grasps in RE and VE. Banana in RE was most commonly grasped using a Medium Wrap [P3] (57.62%, N = 34), followed by Thumb 2-Finger [PC4] (10.16%, N = 6) and Small Diameter [P2] (8.47%, N = 5). Banana in VE was most commonly grasped using a Medium Wrap [P3] (44.82%, N = 26), followed by Small Diameter [P2] (25.86%, N = 15) and Large Diameter [P1] (10.34%, N = 6). Mug in RE was most commonly grasped using a Precision Sphere [PC9] (71.18%, N = 42), followed by Power Sphere [P6] (16.94%, N = 10) and Large Diameter [P1] (6.77%, N = 4). Mug in VE was most commonly grasped using a Large Diameter [P1] (67.24%, N = 39), followed by Small Diameter [P2] (18.64%, N = 11) and Thumb 2-Finger [PC4] (5.08%, N = 3). Lego in RE was most commonly grasped using a Thumb 2-Finger [PC4] (48.33%, N = 29) followed by Thumb-4 Finger [PC6] (20%, N = 12) and Power Sphere [P6] (8.33%, N = 5). Lego in VE was predominantly grasped using Power Sphere [P6] (38.33%, N = 23), followed by Precision Sphere [PC9] (25%, N = 15) and Thumb-Index Finger [PC1] (15%, N = 9). Marker in RE was most commonly grasped using a Thumb-3 Finger [PC5] (54.38%, N =

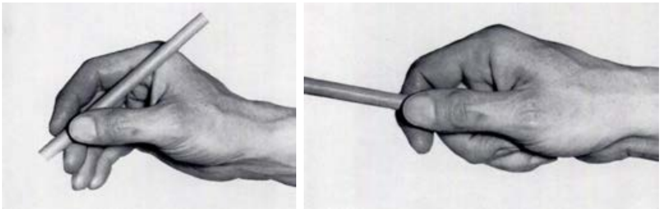
Object	Main Grasps in RE			Main Grasps in VE		
	 [P3] *57.62%*	[PC4] *10.16%*	[P2] *8.47%*	 [P3] *44.82%*	[P2] *25.86%*	[P1] *10.34%*
	 [PC9] *71.18%*	[P6] *16.94%*	[P1] *6.77%*	 [P1] *67.24%*	[P2] *18.64%*	[PC4] *5.08%*
	 [PC4] *48.33%*	[PC6] *20%*	[P6] *8.33%*	 [P6] *38.33%*	[PC9] *25%*	[PC1] *15%*
	 [PC5] *54.38%*	[PC4] *28.07%*	[PC1] *10.52%*	 [P2] *61.66%*	[PC6] *15%*	[PC1] *11.6%*
	 [P3] *55.93%*	[PC5] *16.94%*	[PC4] *6.77%*	 [P1] *63.33%*	[P6] *15%*	[P3] *8.33%*
	 [PC4] *48.33%*	[PC10] *26.66%*	[PC9] *11.66%*	 [PC10] *68.96%*	[PC1] *18.96%*	[P6] *10.34%*
	 [P3] *44.06%*	[P1] *25.42%*	[PC5] *15.51%*	 [P1] *46.55%*	[P6] *20.33%*	[P2] *16.94%*

Table 5.2: Results showing the three most used grasps (with percentages) used in RE condition (Column Main Grasps in RE) and in VE condition (Column Main Grasps in VE) for each individual object used in the study. Each column shows the most used grasps, along with their grasp code detailed in Chapter 4, colour-coded to outline their grasp category: Power grasps in blue and Precision grasps in green.

31), followed by Thumb-2 Finger [PC4] (28.07%, N = 16) and Thumb-Index Finger [PC1] (10.52%, N = 6). Marker in VE was most commonly grasped using a Small Diameter [P2] (61.66%, N = 37), followed by Thumb-4 Finger [PC6] (15%, N = 9) and Thumb-Index Finger [PC1] (11.66%, N = 7). Meat Can in RE was most commonly grasped using a Medium Wrap [P3] (55.93%, N = 33), followed by Thumb-3 Finger [PC5] (16.94%, N = 10) and Thumb-2 Finger [PC4] (6.77%, N = 4). Meat Can in VE was most commonly grasped using a Large Diameter [P1] (63.33%, N = 38), followed by Power Sphere [P6] (15%, N = 9) and Medium Wrap [P3] (8.33%, N = 5). Scissors in RE was most commonly grasped using a Thumb-2 Finger [PC4] (48.33%, N = 29), followed by Precision Disk [PC10] (26.66%, N = 16) and Precision Sphere [PC9] (11.66%, N = 7). Scissors in VE was most commonly grasped using a Precision Disk [PC10] (68.96%, N = 40), followed by Thumb-Index Finger [PC1] (18.96%, N = 11) and Power Sphere [P6] (10.34%, N = 6). Mustard in RE was predominantly grasped using Medium Wrap [P3] (44.06%, N = 26), followed by Large Diameter [P1] (25.42%, N = 15) and Thumb-3 Finger [PC5] (15.51%, N = 9). Mustard in VE was predominantly grasped using Large Diameter [P1] (46.55%, N = 27), followed by Power Sphere [P6] (20.33%, N = 12) and Small Diameter [P2] (16.94%, N = 10).

5.8.2.3 Analysis - Grasp Labels

Results showed a statistically significant difference between grasps used for grasping real objects and grasps chosen when grasping virtual objects. This difference was mainly due to real objects being predominantly grasped using Precision grasps while virtual objects being predominantly grasped using Power grasps. Power grasps are linked to stability and security, being distinguished by large areas of contact between the hand and the object (M. R. Cutkosky, 1989), while Precision grasps present grasps where the object is commonly held between the



(a) Writing position (b) Hammering position

Figure 5.13: Different grasping patterns for different tasks as described by Napier (Napier, 1956).

fingertips to allow an increased level of manipulation. It can be assumed that due to the nature of the simple translate task presented in VR, this increased level of manipulation was not required when interacting with virtual objects, while for real objects users mimicked their real life behaviour. For example, the Marker object was grasped with Precision grasps in all grasping instances collected in RE, users instinctively picking a grasp that allows increased manipulation for writing. This was not the case in VE, where users predominantly grasped the object using a Power grasp. This pattern was identified for other objects used in this study such as Mug, Lego and Scissors. In real grasping, it has been shown that objects are grasped differently based on the purpose of the task. Napier (Napier, 1956) showed that a wooden rod would be grasped between the tip of the thumb and the opposed digits if it was grasped for writing, and grasped between the flexed fingers and the palm if it was grasped for hammering a nail (see Figure 5.13). This observation together with findings in this study introduces the question of whether or not virtual object grasping is influenced by the purpose of the task.

Another important observation is that objects that showed similarities in their shape and size showed similar grasp types. For example, Mug, Meat Can and Mustard are larger objects with bodies that present cylinder form variations and were all predominantly grasped with a Large Diameter [P1]. Napier's (Napier,

1956) theory for real grasping patterns claimed that larger objects are generally grasped with Precision grasps, to provide the greatest span of the hand compatible with stability, however, Feix et al. (Feix et al., 2014b) showed in an observation study that large objects are predominantly grasped with Large Diameter [P1], Sphere-3 Finger [P8] and Medium Wrap [P3] which all belong to Power grasps category. This shows that there might not be a general pattern associated with object size for real objects, but other object characteristics, which supports Griffiths' theory that the hand is primarily conditioned by the shape of the object when performing a stable grasp (Griffiths, 1943). This might also apply for grasping patterns in VR, since the only finding that links grasp types to object size was for larger objects, with medium and smaller objects showing high variability in the grasp types chosen.

5.9 Discussion and Conclusions

Results in this work showed differences in grasping metrics between real and virtual objects, namely for GAp and grasp labels. Participants used apertures larger and smaller than object size in VR, which might be due to the lack of haptic feedback, which has shown to introduce difficulties in estimating virtual object size (Whitwell et al., 2015). This is contrary to real object grasping where haptic feedback guides the hand in matching the object size for a stable grasp. MR literature showed that the lack of haptic feedback led to participants choosing a common GAp, irrespective of object size or shape (Al-Kalbani et al., 2016a) when grasping virtual spheres and cubes. Yet in this experiment, GAp was not consistent across objects, implying that object characteristics did have an influence on grasping patterns, with participants potentially mimicking grasping behaviour from real environments and choosing a different GAp for objects of different sizes. This pattern was also found in grasp label analysis, where different virtual objects showed unique patterns in grasp types chosen for interaction and is consistent with AR gesture elicitation, where it has been found that when asked to design their own gestures, subjects combined a mixture of symbolic and metaphorical gestures reflecting real world interactions (Billinghurst, Piumsomboon, & Bai, 2014).

When comparing grasp labels between RE and VE, significant differences were found, users predominantly grasping real objects using Precision grasps, while grasping the same objects using Power grasps in VR. While the reason for users preferring Power grasps in VR is still to be explored, it can be assumed that due to the nature of VR environments, where there are no real consequences for a failed grasp (such as dropping an object on the floor and breaking it in RE), users did not focus on achieving high precision in their grasps, instead choosing a grasp that is easy to perform and visually correct from their perspective.

While some assumptions about users applying knowledge from real environments to virtual environments hold true at the end of this study, the hypothesis **Grasping patterns for interacting with virtual objects are different than interacting with real objects** is accepted, as differences between RE and VE were shown for GAp and grasp labels. These findings suggest that interaction studies could still use grasping knowledge and taxonomies from real environments to develop virtual interactions, which is the current approach in grasp model development as detailed in Chapter 2. However, this study identified differences and unique grasping patterns that do not occur in real environments, as well as different patterns between unique virtual objects, which shows the need for using virtual grasping taxonomies to fully understand grasping patterns in VR. Yet, a virtual grasp taxonomy has not been developed. To address this, the next chapter presents the first VR grasp taxonomy which provides an overview of grasping patterns for virtual object shape.

6 | Virtual Object Shape

6.1 Introduction

Real grasping literature shows that there is a strong correlation between object shape and grasping patterns (Feix et al., 2014b) (see Chapter 3). However, Chapter 5 showed that there are differences in grasping patterns between real and virtual objects, therefore suggesting that this correlation might not be the same in VR, showing the need for investigating how object shape influences grasping patterns directly in VR. To develop generalisable insights, in real object grasping, researchers followed object categorisation methodologies such as Zingg's (Zingg, 1935) to group objects in meaningful categories that can be connected to grasping patterns (Feix et al., 2014b), and then classified in grasp taxonomies to provide an overall systematic structure which helps researchers reason, compare, elicit and create the appropriate solutions for grasping challenges (Nickerson, Varshney, Muntermann, & Isaac, 2007). Following the methodology of Feix et al. (Feix et al., 2014b) for developing a grasp taxonomy for object characteristics, adapted for virtual environments as detailed in Chapter 4, this chapter aims to explore the influence of object shape on grasping patterns in VR, by categorising 16 virtual objects using Zingg's (Zingg, 1935) methodology and collecting grasping data for these categories. Therefore, 4800 grasp instances were collected in a user elicitation study (N = 50) and labelled and synthesized in the first user-centred VR grasping taxonomy.

The chapter is structured as follows: Section 6.2 presents a literature review of existing work on object categorisation and taxonomies in grasping; Section 6.3 presents the experiment design with a detailed methodology being presented for categorising objects, tasks, hand representation and environment; Section 6.4

presents the protocol for the user elicitation study presented in more detail in Chapter 4; Section 6.5 presents the metrics used for analysing grasping patterns; Section 6.6 presents the proposed hypotheses; Section 6.7 presents the methodology used for data analysis; Section 6.8 presents results where grasp metrics are analysed and structured in the first VR Taxonomy of Grasp Types and Section 6.9 presents discussion and conclusions.

6.2 Background

Due to its dexterity, the human hand is capable of grasping objects of different sizes and shapes, in a manner that can be forceful or delicate depending on object characteristics (Kamakura et al., 1980). For example, the human hand will apply a low force to pick up a pen using a precision grasp, and a higher force to grasp a bottle of water using a power grasp. To classify these observations and facilitate common frameworks of hand usage when interacting with real objects, researchers developed grasp taxonomies that were of high interest in areas such as anthropology (Monaco et al., 2014), hand surgery (Sollerman & Ejeskär, 1995), hand rehabilitation (Lukos et al., 2013) and robotics (Bullock et al., 2013).

Schlesinger (Schlesinger, 1919) took into account object shape and introduced a first simple taxonomy of grasp types, classifying grasping actions and functionality based on different object shapes: cylindrical (for cylindrical objects), tip (for small objects), hook (for heavy objects), palmar (for flat thick objects), lateral (for flat thin objects) and spherical (for spherical objects). A similar approach was followed by Kamakura et al. (Kamakura et al., 1980) who analysed common patterns in finger use for understanding how humans grasp objects based on their shape. For this, they analysed grasping actions of real objects and found that objects that present similar characteristics were grasped using similar patterns, suggesting a direct influence of object characteristics on grasping patterns in real environments. A similar result was found by Landsmeer et al. (Landsmeer, 1962) who showed that the choice of the final position of the grasp is mainly determined by the shape of the object.

To develop grasp taxonomies that provide generalisable insights into grasping patterns based on object characteristics, researchers used object categorisation methods and connected grasp patterns to these categories (for example in the taxonomy of Schlesinger (Schlesinger, 1919), the tip grasp was associated to small objects). Object categorisation dates back to Ancient Greece, when Plato introduced a grouping mechanism based on similar properties (e.g. objects such as apple or olive would be categorised together as spherical objects), and methods for categorisation have been highly used in computer vision (Leibe & Schiele, 2003), for improving text recognition (Kamal & Sultana, 2012), gesture analysis (Hummels & Stappers, 1998) and grasping (Kerzel, Ali, Ng, & Wermter, 2017). A notable categorisation method used for grasping taxonomies is Zingg's (Zingg, 1935) theory that created a framework that uses three primary object dimensions (A, B and C) and categorises objects based on the relationship between dimensions. Feix et al. (Feix et al., 2014b) used this methodology for categorising objects in shape categories and analysing grasping patterns associated with them and found that object types have very specific grasping patterns associated with them, with long objects being usually grasped in a wrap grasp, disk objects being predominantly grasped from the side with a precision disk grasp and small and lightweight objects being grasped using thumb-2 finger grasps.

While real grasping taxonomies showed a strong correlation between object shape and grasping patterns, Chapter 5 identified differences in grasping patterns between real and virtual objects. Therefore, the knowledge provided by these taxonomies cannot be directly applied to virtual grasping without expecting unwanted unknowns to be present in the grasping interaction, showing a need for investigating this correlation between object shape and grasping patterns directly in VR. While Chapter 5 outlined some correlations between objects with similar characteristics and grasping patterns (e.g. smaller objects presented larger grasp apertures and were predominantly grasped using power grasps), this correlation is still unclear and needs to be fully explored. To address this, this chapter aims to explore the correlation between object shape and grasping patterns in VR and present this correlation in a first VR grasping taxonomy that aims to provide a foundation for understanding existing challenges with natural grasping in VR.

6.3 Experiment Design

To analyse the correlation between object shape and grasping patterns in VR, and develop the first VR grasping taxonomy, a user study (N = 50) was conducted on 16 virtual objects from the *"Yale-Carnegie Mellon University-Berkeley Object and Model Set"* (Calli et al., 2015), chosen following the methodology detailed in Chapter 4 and categorised using Zingg's methodology (Zingg, 1935), which has been used in grasping research (Feix et al., 2014b). This section provides a detailed overview of the virtual object categorisation methodology, task, apparatus, environment and 3D hand model used for interaction as detailed in Table 4.3.

6.3.1 Virtual Objects Categorisation

Zingg's (Zingg, 1935) methodology categorises objects based on their shape and the three dimensions that indicate the volume of geometric bodies. Zingg (Zingg, 1935) defined A as the longest dimension of an object, C the shortest and B the remaining dimension. He defined a constant R to describe the relationship between dimensions and categorise the object, determining that the value at which

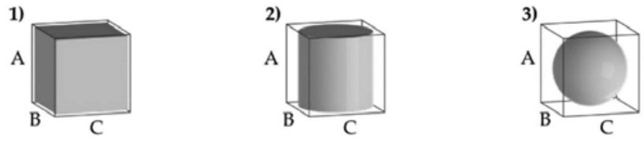
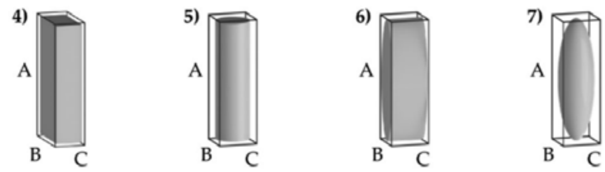
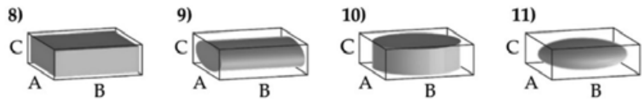
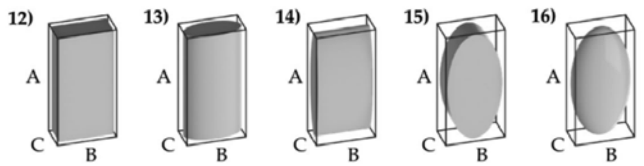
Zingg's	Definition	Example Objects
Equant	$B \leq A < RB$ $C \leq B < RC$	
Prolate	$A > RB$ $C \leq B < RC$	
Oblate	$B \leq A < RB$ $B > RC$	
Bladed	$A > RB$ $B > RC$	

Table 6.1: Object categories defined by Zingg (Zingg, 1935) based on object dimensions A, the longest dimension, C, the shortest dimension and B the remaining dimension. Column "Zingg's" presents the four categories proposed. Column "Definition" presents the mathematical expressions that represent the relationship between object dimensions (A, B and C) for each category. Column "Example Objects" shows example objects for each category, used for reference in grasping literature (Feix et al., 2014b).

one typically regards two axes to be different is about R = 3/2 (Zingg, 1935). Based on these parameters, four shape categories were defined as part of Zingg's categorisation framework: *Equant*, *Prolate*, *Oblate* and *Bladed*. Table 6.1 shows the four categories with their definition and example objects. Each category is defined by two mathematical expressions that need to be true for an object with dimensions A, B and C to be categorised within that category. As an example, Figure 6.1 shows the categorisation of the banana object, which has the following dimensions: A = 190 mm, B = 36 mm and C = 36 mm. If the mathematical ex-

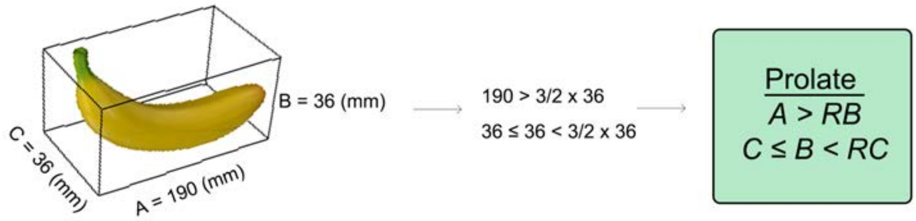


Figure 6.1: Categorisation of Banana object. Object dimensions are used to verify the mathematical expressions of each of the Zingg's (Zingg, 1935) categories. With A = 190, B = 36 and C = 36, and the constant R = 3/2, the Prolate's mathematical expressions are verified, therefore categorising the banana object in the Prolate category.

pressions for all categories are verified, it can be observed that the expressions are true for the Prolate category ($A > RB, C \leq B < RC$) therefore categorising the banana object as Prolate.

The virtual objects used for this experiment were the 16 objects selected from the *Yale-Carnegie Mellon University-Berkeley Object and Model Set* (Calli et al., 2015) following the methodology detailed in Chapter 4. These objects were then grouped in Zingg's (Zingg, 1935) shape categories following the methodology presented above. Table 6.2 shows the A, B and C dimensions of each object and their categorisation in the four shape categories.

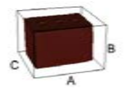
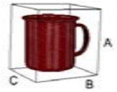
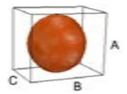
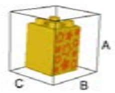
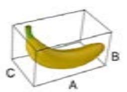
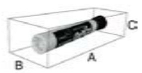
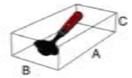
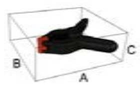
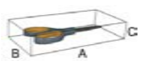
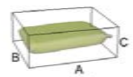
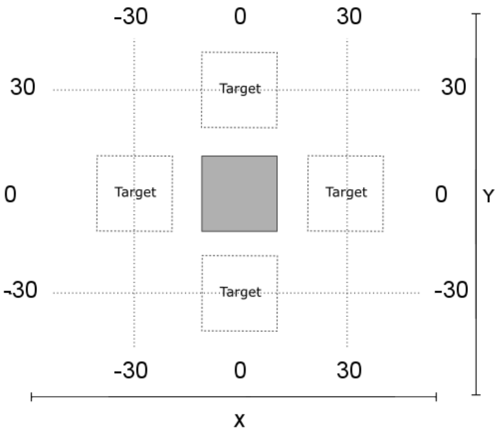
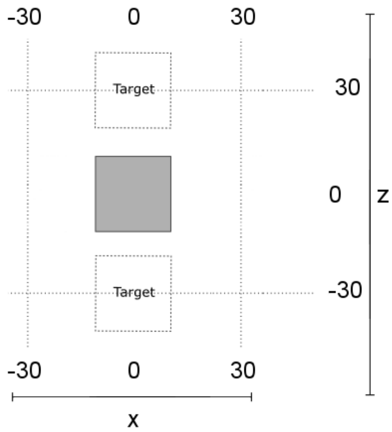
Zingg	Object	A(mm)	B(mm)	C(mm)	Object	A(mm)	B(mm)	C(mm)
Equant $B \leq A < RB$ $C \leq B < RC$		75	50	50		82	80	80
		73	73	73		26	26	26
Prolate $A > RB$ $C \leq B < RC$		190	36	36		135	32	24
		121	18	18		195	20	14
Oblate $B \leq A < RB$ $B > RC$		210	158	60		115	90	27
		85	73	28		97	82	50
Bladed $A > RB$ $B > RC$		190	95	58		200	87	14
		250	98	65		114	72	14

Table 6.2: Virtual objects used in this experiment categorised in Zingg's shape categories (Zingg, 1935): Equant, Prolate, Oblate and Bladed based on their dimensions (A, B and C).



(a)



(b)

Figure 6.2: Spatial distribution of targets for the translation tasks. (a) shows target distribution for translation tasks in $\pm X$ and $\pm Y$. (b) displays targets for translation in $\pm Z$. Axes in centimetres

6.3.2 Task

Following on from the simple translate tasks in Chapter 5, the tasks selected for this experiment were again a set of simple translate tasks, which are the most common interaction tasks in virtual environments (D. Chen et al., 2018; Hartney et al., 2019). The tasks were defined in the three axes of the Cartesian coordinate system (*X*, *Y*, *Z*), in both positive and negative directions. Participants were asked to move each virtual object to a target position, 30 cm away in each direction. A representation of the six different translate tasks is shown in Figure 6.2.

6.3.3 Apparatus

As in the virtual environment presented in Chapter 5, a custom experimental framework was built using the Oculus DK2 VR headset, the Leap Motion device and a Logitech Pro 1080p HD. As detailed in Chapter 4 (Section 4.5.1), the webcam was attached on the HMD to capture participants' hands at all times during the experiment. Pilot tests were conducted to find the optimal position for attaching the camera to facilitate recording of participants' hands at all times. The Leap Motion Controller was attached to the HMD, facing the user's hands. The starting position was the same for each participant.

6.3.4 Hand Representation

As in Chapter 5, the 3D hand model used for interaction was the abstract hand, which was extracted from the Leap Motion SDK and represented as a set of cylinders and spheres representing bones and joints. This consistency in hand representation between the two experiments allows an accurate comparison between grasp metrics recorded in the two experiments.

6.3.5 Environment

Aiming to replicate a similar virtual environment as Chapter 5, the user experiment was conducted in a controlled environment under laboratory conditions. The test room was lit by a 2700k (*warm white*) fluorescent with no external light source. The virtual environment was composed of a virtual table and a virtual shelf as shown in Figure 6.3. For each task, a virtual object (one of the 16 chosen objects, in randomised order) would appear on the virtual table, together with the marker for the target position as detailed in Chapter 4. The shelf was used for creating a realistic context for translate $\pm Y$ tasks.

6.3.6 Participants

A total of 50 right-handed participants (23 females, 27 males) from a population of university students and staff members volunteered to take part in this study. Participants ranged in age from 19 to 65 (M = 29.4, SD = 12.45). All partici-

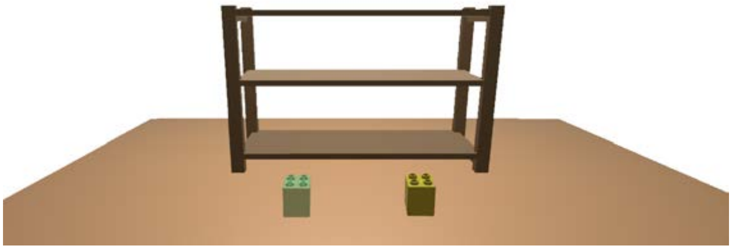


Figure 6.3: Interaction environment displaying a virtual table, the virtual object (Lego), the marker for the target (semi-transparent green Lego) and the virtual shelf.

pants performed the 6 experiment tasks with the 16 virtual objects. Participants completed a standardised consent form and were not compensated. Visual acuity of participants was measured using a Snellen chart, each participant was also required to pass an Ishihara test to check for colour blindness. As with Chapter 5, participants with colour blindness and/or non corrected visual acuity of $<$ 0.80 (where 20/20 is 1.0) were not included in this study. Participants were asked to self-assess their level of experience with VR systems, with 19 participants reporting to have an average level of experience, 25 reported being novice to the technology and 6 self-labelled themselves as experts. No participant had any significant experience with hand tracking sensors.

6.4 Protocol

6.4.1 Pre-test

Prior to the study, participants were given a written informed consent where the test protocol and main aim of the study was described. Additionally, participants completed a pre-test questionnaire enquiring about their background level of experience with VR systems and hand recognition sensors.

6.4.2 Training

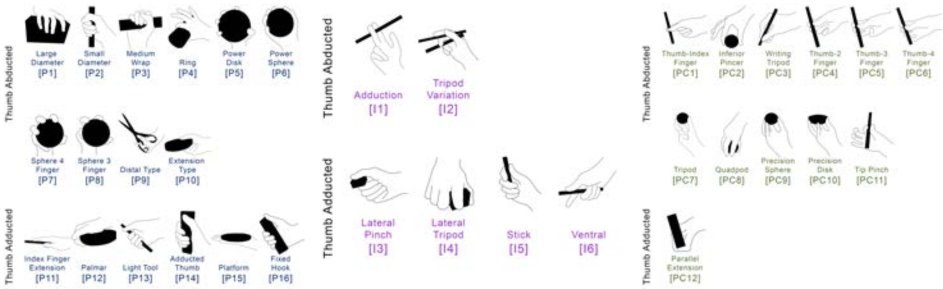
Participants underwent initial hand interaction and task training to familiarise themselves with the VR environment and hand interaction space. This training task was a representative version of the tasks in the user study, where they were asked to grasp and translate a cube object (as in Chapter 5) in the 3D space while being seated.

6.4.3 Test

Once participants were comfortable with the interaction space and the overall VR environment, they were presented with the main experimental task. Participants were seated during the experiment. Each participant completed 96 grasps (16 objects $\times$ 6 tasks), with a total of 4800 grasps recorded during the study (96 grasps $\times$ 50 participants). Objects and tasks were presented in randomised order. As with Chapter 5, a Wizard of Oz approach was followed and participants were instructed to grasp the virtual objects the way they felt most intuitive, notifying the test instructor when they were happy with their grasp.

6.4.4 Post-test

After all tasks were completed, participants were asked to complete a post-test questionnaire comprised of the NASA-TLX questionnaire (Hart & Staveland, 1988) and the Motion-Sickness Questionnaire (MSAQ) (J. Gianaros, Muth, Mordkoff, Levine, & M. Stern, 2001).



(a) Power Grasps (b) Intermediate Grasps (c) Precision Grasps

Figure 6.4: Grasp types from the real grasp taxonomy (Feix et al., 2009)

6.5 Metrics

Post-test questionnaire: To ensure that there is no undue bias in the methodology for data collection due to different cognitive load or motion sickness in participants, this experiment reports on NASA-TLX (Hart & Staveland, 1988) for evaluating perceived user workload during the task and Motion Sickness Questionnaire (MSAQ) to assess motion sickness (J. Gianaros et al., 2001) on a scale from 0 to 100.

Grasp Aperture (GAp): GAp is a metric used to measure how accurately users estimate the size of a virtual object and has been described in more detail in Chapter 4.

Grasp Labels: As in Chapter 5, the grasp categories are Power, Intermediate and Precision. Power grasps are linked to stability and security (Figure 6.4 a). Intermediate grasps present elements of Power and Precision roughly in the same proportion, enabling a finer representation of grasp types (Bullock et al., 2013) (Figure 6.4 b). Precision grasps present grasps where the object is commonly held between the fingertips (Figure 6.4 c).

User Grasp Choice Agreement: Chapter 5 showed that some objects presented more variability between participants in grasping patterns than others. Therefore, in this experiment, user grasp choice agreement is analysed for each object to understand if there is a link between object shape and grasp variability. The grasp agreement score was defined as the agreement among the grasp types proposed by participants per object, following the definition of Wobbrock et al. (Wobbrock et al., 2005) and was computed using the equation:

$$\frac{\sum_{r \epsilon R} \sum_{P_i \subseteq P_r} \left(\frac{P_i}{P_r}\right)^2}{|R|} \tag{6.1}$$

Where *r* is a referent in the set of all referents available for each object or task (segmented by object category) *R*; P_r is the set of grasp proposals for referent *r* and P_i is a subset of identical grasp labels for P_r as in (Wobbrock et al., 2005, 2009).

6.6 Hypotheses

Chapter 5 showed that while differences exist between grasping real and virtual objects, there are also similarities, with results suggesting there is a link between object shape and grasping patterns, which has been shown in immersive virtual object grasping (Al-Kalbani et al., 2016a). Therefore, the following hypothesis is proposed:

H_1**:** Virtual object shape has an effect on grasping patterns in VR.

To understand if grasping patterns change for different translate tasks, which is currently unknown in VR, the following null hypothesis is proposed:

H_2**:** Translate tasks do not have an effect on grasping patterns in VR.

6.7 Data Analysis

The Shapiro-Wilk (Shapiro & Wilk, 1965) normality test found the data to be not normally distributed. Data collected for GAp was non-parametric and not normally distributed, therefore statistical significance between the four dependent groups (Equant, Prolate, Oblate and Bladed) where the variable of interest is continuous (GAp in mm) was tested using the Friedman test (Friedman, 1940) with an alpha of 5%. For testing statistical significance between grasp patterns for object categories, where the dependent variable (Grasp Category) is nominal categorical (Power, Intermediate and Precision), contingency tables were created and analysed for significance using a Chi-Squared Test of Independence with 95% Confidence Intervals, therefore, a p-value of less than .05 will indicate statistical significance. Cramer's V calculation for effect sizes was applied after verifying its assumptions that the variables under analysis are categorical. Results were interpreted following existing guidelines based on degrees of freedom.

6.8 Results

6.8.1 NASA-TLX and MSAQ

No participants reported motion sickness before and during the experiment. The mean MSAQ score reported is 24.69 (SD = 14.98) when the maximum score is 100. The mean score for gastrointestinal items was 21.88 (SD = 18.82). The mean score for central items was 28.13 (SD = 20.43). The mean score for peripheral items was 22.07 (SD = 15.97). The mean score for sopite-related was 25.16 (SD = 17.55). This shows there was a negligible effect on motion sickness during this experiment.

The NASA-TLX score was 28.89 (SD = 16.78), which based on existing guidelines is considered a medium workload. These results show that the tasks did not suppose a challenge for participants and they did not feel overloaded.

6.8.2 Grasp Aperture (GAp)

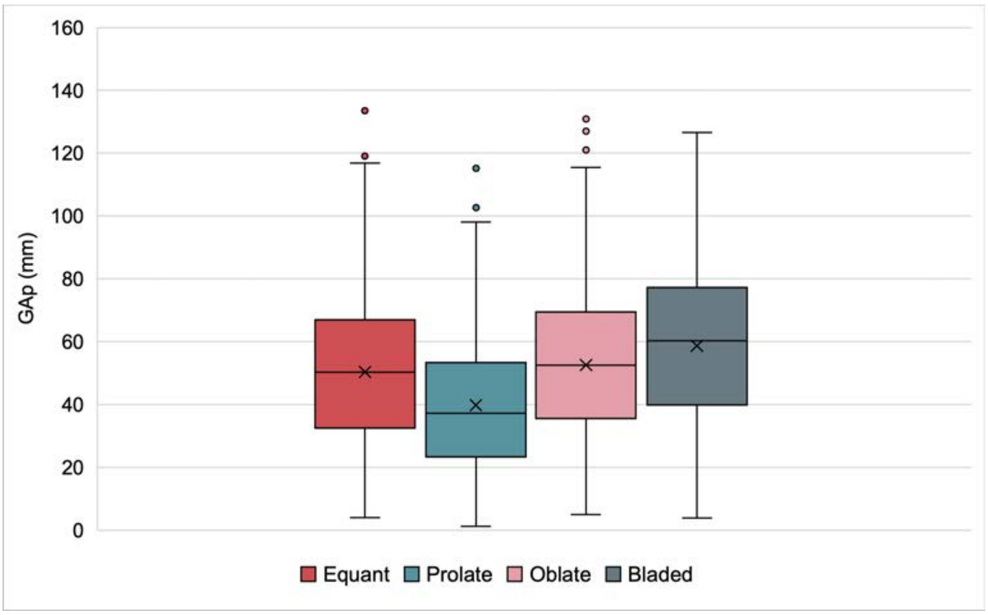


Figure 6.5: GAp in mm for virtual objects categorised based on Zingg's (Zingg, 1935) methodology. X marks on boxplots indicate the mean GAp across all participants for Equant, Prolate, Oblate and Bladed. The line that divides the box in two plots shows the median, which marks the mid-point of the data. Whiskers represent the highest and lowest values within 1.5 times the interquartile range. Outliers are shown in coloured circles.

Figure 6.5 shows an overview of GAp for each object category presented in this experiment (Equant, Prolate, Oblate and Bladed). Differences in GAp between categories can be observed as follows: Equant (M = 50.27, SD = 23.66), Prolate (M = 37.21, SD = 20.55), Oblate (M = 52.45, SD = 24.09) and Bladed (M = 60.25, SD = 25.35). However, a non-parametric Friedman test of differences showed that these differences in GAp between object categories were not significant χ^2 *(3, N*

= 4717) = 146.21, p = 1.726.

Figures 6.6, 6.7, 6.8, 6.9 show GAp for individual objects within the same category, with A, B,C points plotted to represent object size on each dimension in mm (see Table 6.2 for A, B, C values). Figure 6.6 shows GAp(in mm) for objects in Equant category: Brick (M = 55.47, SD = 22.77), Mug (M = 51.56 SD = 31.33), Orange (M = 55.07, SD = 18,82) and Lego (M = 40.5, SD = 16.14) Figure 6.7 shows GAp (in mm) for objects in Prolate category: Banana (M = 49.55, SD = 20.02), Hammer (M = 35.90, SD = 20.85), Marker (M = 30.98, SD = 19.31) and Spoon (M = 36.92, SD = 20.26). Figure 6.8 shows GAp (in mm) for objects in Oblate category: Crackers Box (M = 66.46, SD = 27.79), Clamp (M = 45.36, SD = 20.92), Gelatine Box (M = 44.39, SD = 17.57) and Meat Can (M = 63.91, SD = 23.23). Figure 6.9 shows GAp (in mm) for objects in Bladed category: Mustard (M = 59.82, SD = 25.33), Scissors (M = 49.64, SD = 24.83), Cleanser Bottle (M = 71.23, SD = 25.65) and Sponge (M = 60.25, SD = 22.10).

6.8.2.1 Analysis - GAp

Results of this experiment showed that differences in GAp between object shape categories were not significant, therefore suggesting that in VR, GAp is not directly influenced by object shape if objects are categorised using the dimension representation from Zingg's method (Zingg, 1935). In real grasping, grasp aperture is influenced by object size. This effect was observed in Chapter 5 and can be observed in the results of this study, with Prolate category, which presents long and narrow objects showing the smallest GAp (see Figure 6.5) when compared to the other object categories. However, to fully understand how GAp is linked to object size, Figures 6.6, 6.7, 6.8 and 6.9 show GAp for individual objects with markers for object size as a reference.

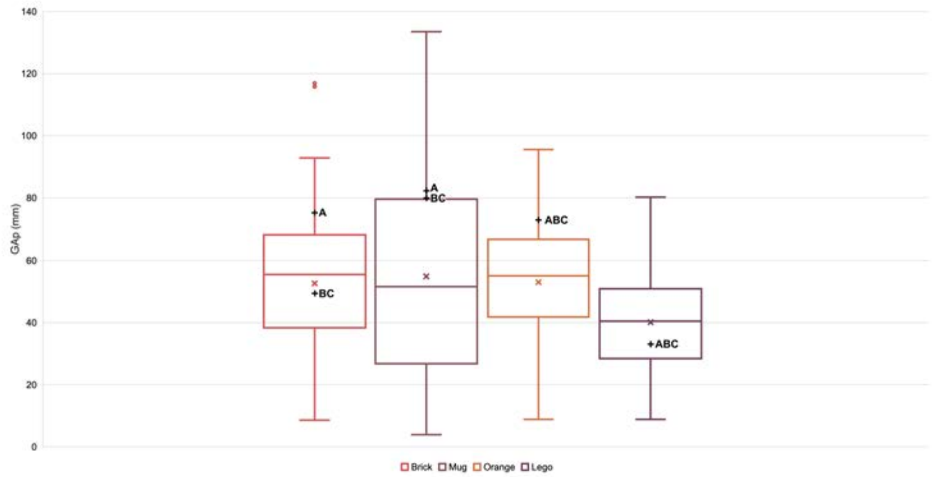


Figure 6.6: GAp in mm for individual objects within Equant category. X marks on boxplots indicate the mean GAp across all participants. The line that divides the box in two plots shows the median, which marks the mid-point of the data. Whiskers represent the highest and lowest values within 1.5 times the interquartile range. Outliers are shown in circles. A, B and C marks represent virtual object dimensions, as detailed in Table 6.2 plotted as a reference for how GAp relates to individual object sizes.

Equant objects presented the highest variability in terms of GAp patterns, with spherical/cylindrical objects being grasped smaller than all object dimensions, such as Mug and Orange, while cuboid objects being grasped smaller and larger than object dimensions (Brick and Lego) as shown in Figure 6.6. This is consistent with MR grasping research, where it has been shown that participants grasped spherical objects smaller than cuboid objects (Al-Kalbani et al., 2016a). However, it should be noted that within this category there was a high variability in the overall object size, which therefore might have influenced the GAp.

Objects in Prolate category showed more similarities in terms of GAp, all objects being predominantly grasped larger than dimensions B and C and always smaller than dimension A as shown in Figure 6.7. This might have been due to the nature of Prolate objects which are characterised by long and narrow bodies, where di-

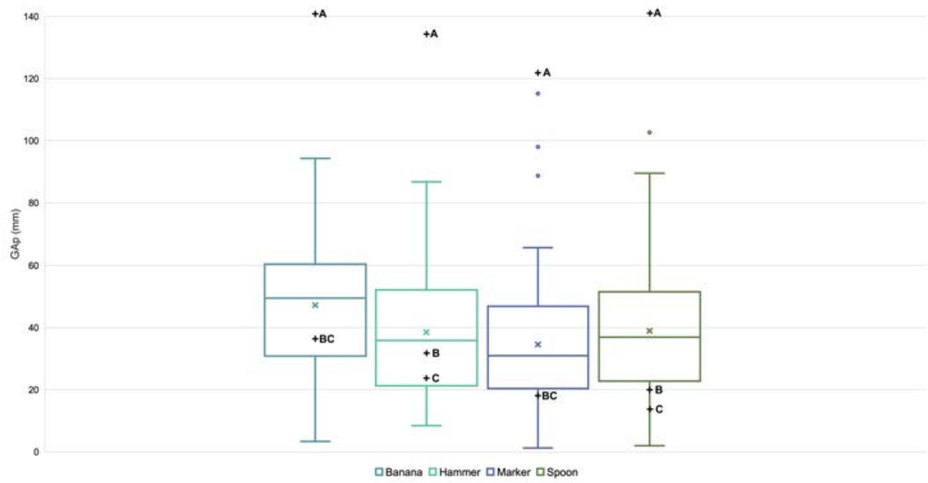


Figure 6.7: GAp in mm for individual objects within Prolate category. X marks on boxplots indicate the mean GAp across all participants. The line that divides the box in two plots shows the median, which marks the mid-point of the data. Whiskers represent the highest and lowest values within 1.5 times the interquartile range. Outliers are shown in circles. A, B and C marks represent virtual object dimensions, as detailed in Table 6.2 plotted as a reference for how GAp relates to individual object sizes.

mension A is significantly bigger than dimension B and C. For example, objects used in this experiment have A larger than 120 mm, considering that in real environments comfortable grasps are typically less than 70 mm (Feix et al., 2014b), while MR research showing that the most comfortable grasp size is 80 mm (Al-Kalbani et al., 2016a), therefore suggesting that grasping along this dimension would not be intuitive for the users. A similar pattern was found for objects in Oblate and Bladed categories (Figures 6.8 and 6.9) which were predominantly grasped smaller than dimension A and B, but larger than dimension C.

Real object grasping research showed that the human hand has a tendency to grasp the smallest dimension of an object (C). If this were the case for virtual objects, this would mean users overestimated object size for all object categories, with some exceptions in Equant category as discussed above. However, since Chap-

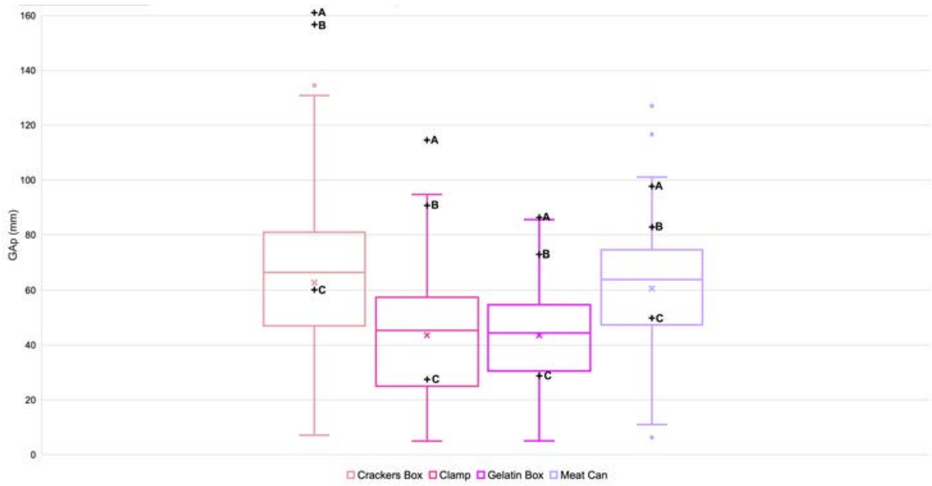


Figure 6.8: GAp in mm for individual objects within Oblate category. X marks on boxplots indicate the mean GAp across all participants. The line that divides the box in two plots shows the median, which marks the mid-point of the data. Whiskers represent the highest and lowest values within 1.5 times the interquartile range. Outliers are shown in circles. A, B and C marks represent virtual object dimensions, as detailed in Table 6.2 plotted as a reference for how GAp relates to individual object sizes.

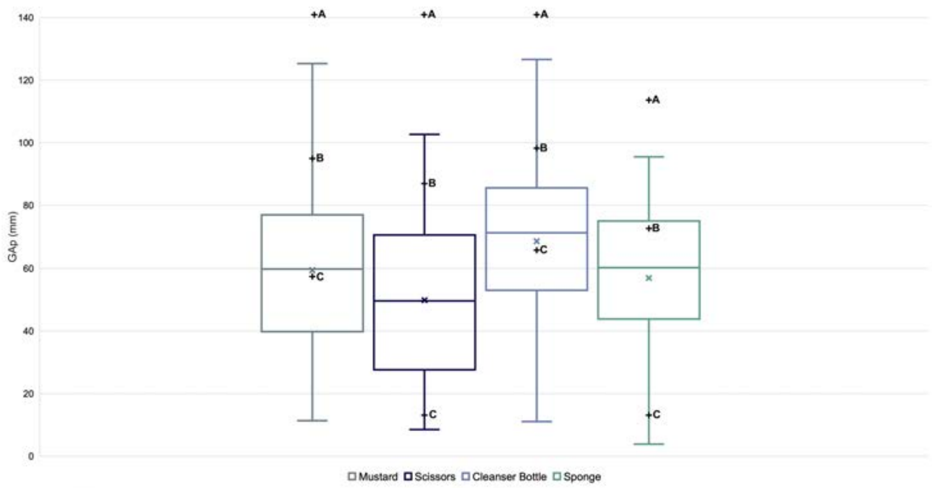
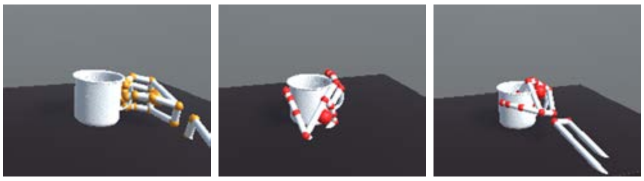


Figure 6.9: GAp in mm for individual objects within Bladed category. X marks on boxplots indicate the mean GAp across all participants. Whiskers represent the highest and lowest values within 1.5 times the interquartile range. Outliers are shown in circles. A, B and C marks represent virtual object dimensions, as detailed in Table 6.2 plotted as a reference for how GAp relates to individual object sizes.



(a) Mug Grasped Location - Handle (b) Mug Grasped Location - Body (c) Mug Grasped Location - Top

Figure 6.10: Examples of multiple locations for grasping a virtual object.

ter 5 showed differences in user grasp choice for real and virtual objects, there is also the possibility that users grasped along dimension A or B, underestimating the object size due to the lack of haptic feedback (Murcia-López & Steed, 2018). Yet, the validity of these assumptions is highly dependent on where users grasped the virtual objects. For example, virtual objects that have handles (Mug, Scissors and Clamp) and therefore provide more grasping possibilities, showed a higher variability in GAp, which might be due to users grasping them on different locations, as shown in Figure 6.10. As discussed in Chapter 5, assuming that users aim to mimic real grasping in VR, GAp would then be influenced by the size of the grasped location, showing the need for analysing grasped location together with GAp to fully understand grasping behaviour in VR, which will be presented in more detail in Chapter 7.

6.8.3 Grasp Labels

A total of 4800 grasps were recorded during the experiments (50 participants x 16 objects x 6 tasks) which were labelled following the methodology presented in Chapter 4. Out of 4800 grasps, 42 were removed due to being rated as "Cannot Classify" by at least one of the raters and 41 were removed due to disagreement between virtual and real view caused by sensor errors. The remaining 4717 (1159 for Equant objects, 1183 for Prolate objects, 1187 for Oblate objects and 1188

Object	X			Y			Z			Statistic
	P	PC	I	P	PC	I	P	PC	I	
Equant	76%	22%	1%	79%	20%	0.5%	77%	21%	0.25%	*Stat = 668.91, p = .261*
Prolate	67%,	16%	16%	65%	21%	13%	66%	18%	14%	*Stat = 3.81, p = .432*
Oblate	75%	23%	0.76%	75%	24%	0.5%	75%	24%	0.25%	*Stat = 1.08, p = .896*
Bladed	53%	45%	1.51%	55%	44%	0.50%	51%	47%	0.25%	*Stat = 5.66, p = .225*

Table 6.3: Grasp category results for object categories and tasks presented in this study. Percentages for use of each grasp category (P for Power, PC for Precision and I for Intermediate) are shown for each object category, together with the statistical results for comparing between task conditions.

for Bladed objects) were further analysed for developing the first VR grasping taxonomy.

Cohen's Kappa was used to measure inter-rater reliability for labelling the grasps. Raters agreed in 86% of instances (Cohen's Kappa = 0.4) which based on existing guidelines is a moderate agreement, which is often achieved when subjectivity is involved in the process (S. Sun, 2011) and is a common agreement score for classification tasks (Feix et al., 2014b).

6.8.3.1 Grasp Category

Power grasps were the most used grasps in this experiment (67.39%, N = 3179) of the total dataset, with the remaining instances being Precision grasps (14.54%, N = 686) and Intermediate grasps (3.24%, N = 153) as shown in Figure 6.11.

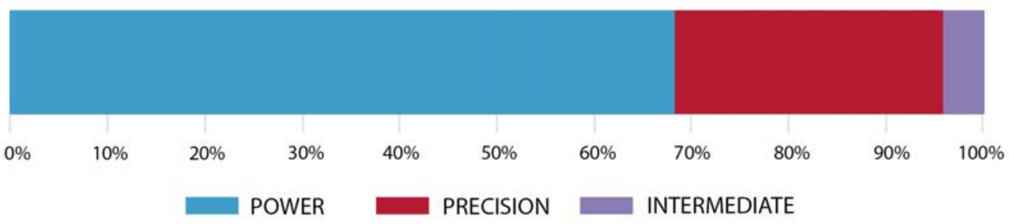


Figure 6.11: Use of Power, Precision and Intermediate grasps in this user experiment. Power grasps were the most used grasp types (67.39%, N = 3179) followed by Precision grasps (14.54%, N = 686) and Intermediate grasp types (3.24%, N = 153).

Task: To understand if task influenced grasp choice, grasp patterns for every task (Translate X, Translate Y and Translate Z) were compared for every category of virtual objects (Equant, Prolate, Oblate and Bladed). Equant objects for Translate X were predominantly grasped using Power grasps (76.37%, N = 299), followed by Precision grasps (22.19%, N = 87) and Intermediate grasps (1.53%, N = 6); For Translate Y they were predominantly grasped using Power grasps (79.09%, N = 314) followed by Precision grasps (20.40%, N = 81) and Intermediate grasps (0.5%, N = 2); For Translate Z they were predominantly grasped using Power grasps (77.86%, N = 306), followed by Precision grasps (21.88%, N = 86) and Intermediate grasps (0.25%, N = 1). A Chi-Squared test of Independence showed that this difference was not statistically significant *(χ^2 (2, N=1182) = 668.91, p = .261).*

Prolate objects for Translate X were predominantly grasped using Power grasps (67.01%, N = 256), followed by Precision grasps (16.49%, N = 63) and Intermediate grasps (16.49%, N = 63); For Translate Y they were predominantly grasped using Power grasps (65.63%, N = 254), followed by Precision grasps (21.18%, N = 82) and Intermediate grasps (13.17%, N = 51); For Translate Z they were predominantly grasped using Power grasps (66.83%, N = 262), followed by Precision grasps (18.87%, N = 74) and Intermediate grasps (14.28%, N = 56). A Chi-Squared test of Independence showed that this difference was not statistically significant *(χ^2 (2, N=1161) = 3.81, p = .432).*

Oblate objects for Translate X were predominantly grasped using Power grasps (75.57%, N = 297), followed by Precision grasps (23.66%, N = 93) and Intermediate grasps (0.76%, N = 3). For Translate Y they were predominantly grasped using Power grasps (75%, N = 297), followed by Precision grasps (24.49%, N = 97) and Intermediate grasps (0.5%, N = 2); For Translate Z they were predominantly grasped using Power grasps (75.50%, N = 299), followed by Precision grasps (24.24%, N = 96) and Intermediate grasps (0.25%, N = 1). A Chi-Squared test of Independence showed that this difference was not statistically significant *(χ^2 (2, N=1185) = 1.08, p = .896)*.

Bladed objects for Translate X were predominantly grasped using Power grasps (53.28%, N = 211), followed by Precision grasps (45.20%, N = 179) and Intermediate grasps (1.51%, N = 6); For Translate Y they were predominantly grasped using Power grasps (55.18%, N = 218), followed by Precision grasps (44.30%, N = 175) and Intermediate grasps (0.50%, N = 2); For Translate Z they were predominantly grasped using Power grasps (51.88%, N = 206), followed by Precision grasps (47.85%, N = 190) and Intermediate grasps (0.25%, N = 1). A Chi-Squared test of Independence showed that this difference was not statistically significant *(χ^2 (2, N=1188) = 5.66 p = .225)*. A visual representation of these results can be seen in Figure 6.12.

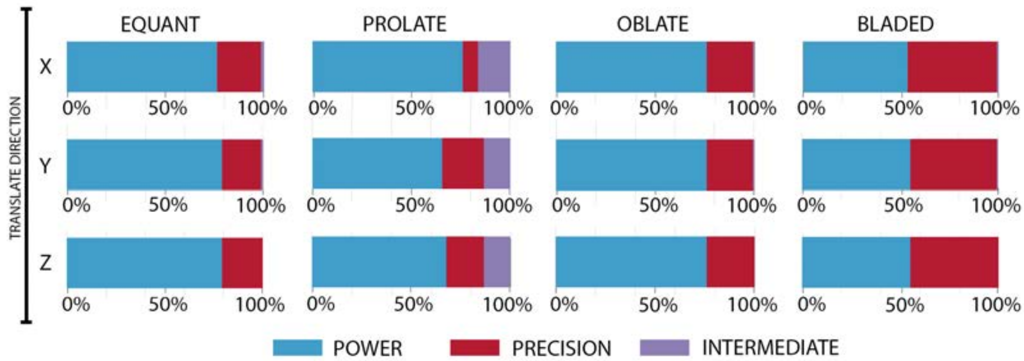


Figure 6.12: Grasp categories (Power, Intermediate and Precision) used for virtual object categories (Equant, Prolate, Oblate and Bladed) presented for each task (Translate X, Translate Y and Translate Z).

Object Category: When comparing grasp labels between object categories (Equant, Prolate, Oblate and Bladed) there were differences in grasp category (Power, Intermediate and Precision). Equant objects were predominantly grasped using Power grasps (77.48%, N = 898), followed by Precision grasps (21.82%, N = 253) and

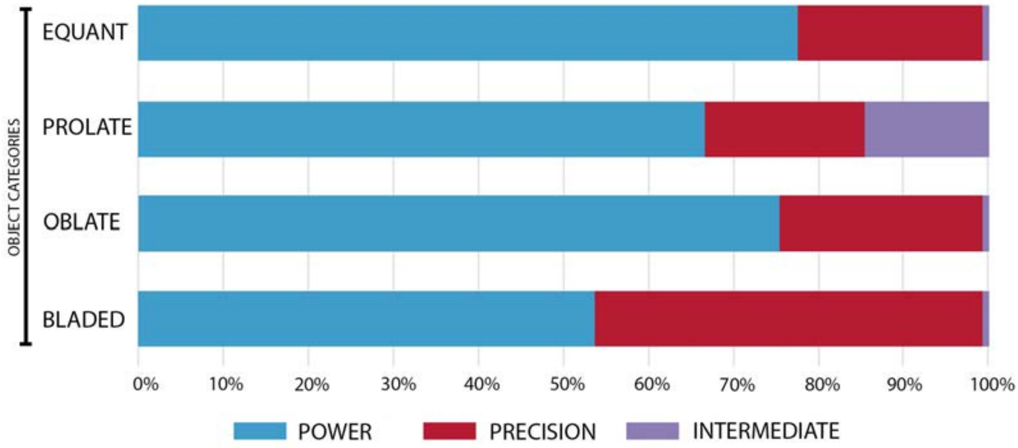


Figure 6.13: Grasp categories (Power, Intermediate and Precision) used for virtual object categories (Equant, Prolate, Oblate and Bladed).

Intermediate grasps (0.69%, N = 8); Prolate objects were predominantly grasped using Power grasps (67.11%, N = 794), followed by Precision grasps (18.51%, N = 219) and Intermediate grasps (14.37%, N = 170); Oblate objects were predominantly grasped using Power grasps (75.23%, N = 893), followed by Precision grasps (24.34%, N = 289) and Intermediate grasps (0.42%, N = 5); Bladed objects were predominantly grasped using Power grasps (53.45%, N = 635) followed by Precision grasps (45.79%, N = 544) and Intermediate grasps (0.75%, N = 9). A Chi-Squared test of Independence showed that this difference was statistically significant (**χ^2 (3, N=4717) = 668.91, p < .001***) with a medium ES (Cramer's V = 0.26). A visual representation of these results can be seen in Figure 6.13.

6.8.3.2 Most Common Grasps

Figure 6.14 shows the most common grasp types and their usage percentages. Large Diameter [P1] from the Power grasp category was the most prevalent grasp type, accounting for 38.75% of the labels in the data set (N = 1828). This grasp was followed by the Precision Disk [PC10] grasp from the Precision grasp cate-

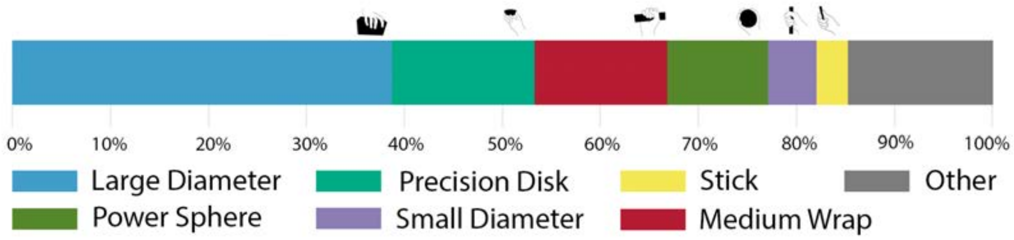


Figure 6.14: Most common used grasp types in this experiment. The six most used grasp types accounted for more than 85% of the labelled data, with the most used grasp type being Large Diameter [P1].

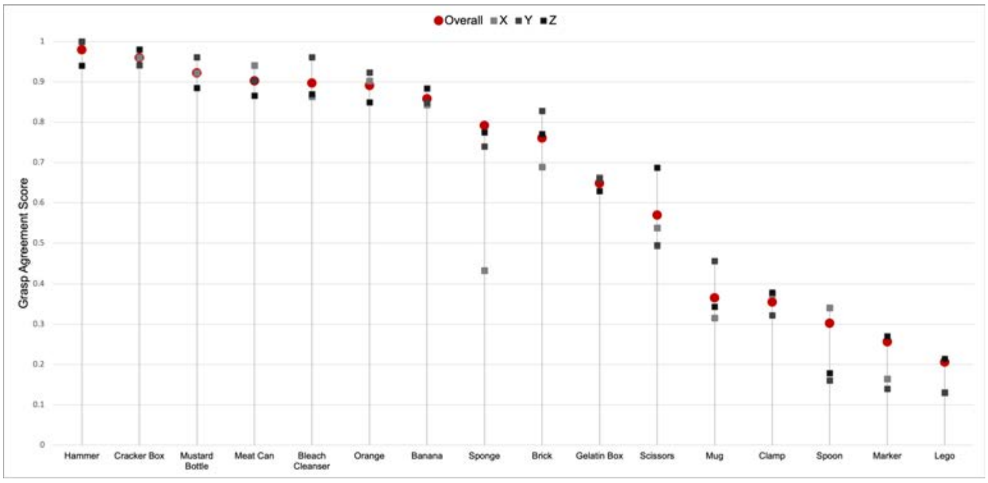


Figure 6.15: Agreement on grasp choice between participants, showing a notable group of objects presenting a high agreement score (≥ 0.90).

gory with 14.54% of the labels (N = 686) and the Medium Wrap [P3] from the Power grasp category with 13.44% of the labels (N = 634). In total, the six most used grasps accounted for 85.18% (N = 4018) of the labelled data.

6.8.3.3 User Grasp Choice Agreement

Agreement on grasp choice between participants is presented in Figure 6.15. Agreement scores should be 100% when all the proposed grasps are identical and $\approx 0\%$ when they are unique. Hammer showed an agreement of 97.98%; Crackers Box showed an agreement of 96.01%; Mustard showed an agreement of 92.22%; Meat can showed an agreement of 90.23%; Cleanser bottle showed an agreement of 89.71%; Orange showed an agreement of 89.07%; Banana showed an agreement of 85.80%; Sponge showed an agreement of 79.15%; Brick showed an agreement of 76.04%; Gelatine Box showed an agreement of 64.83%; Scissors showed an agreement of 56.92%; Mug showed an agreement of 36.53%; Clamp showed an agreement of 35.52%; Spoon showed an agreement of 30.24%; Marker showed an agreement of 25.64%; Lego showed an agreement of 20.58%.

When looking at object shape categories, Equant objects showed an overall agreement of 55.56%, Prolate objects showed an agreement of 59.92%, Oblate objects showed an agreement of 71.65% and Bladed objects showed an agreement of 79.50%.

6.8.3.4 Analysis - Grasp Labels

Results showed that grasp choice was not influenced by simple translate tasks while being primarily influenced by virtual object shape. While real grasping literature showed that grasp choice is highly influenced by the intended task (Napier, 1956) as discussed in Chapter 5, these results do not invalidate the possibility of this being the case in VR, since this experiment only explored translate tasks in different directions. In MR grasping literature it has been shown that occlusion between the user's hands and parts of the virtual object influenced participants' grasping on the Z axis, however in this experiment no significant differences in grasp patterns were found between translation on X, Y and Z axis.

However, differences were found between Equant, Prolate, Oblate and Bladed, showing that users grasped objects in a similar manner regardless of the task, changing their patterns for objects of different shape. Oblate and Equant objects presented a similar grasp category segmentation, with more than 70% of the grasp choices being from the Power grasp category, while the remaining choices were from Precision category. The use of Intermediate grasps was negligible (below 1%). The same negligible pattern for Intermediate grasps was found in the Bladed category, however, with these objects the split between Precision and Power grasps is circa 50%. For Prolate objects, due to their distinctive tubular shape, the use of Intermediate grasps increased to 15%, showing a similar pattern to real grasping approach, where Intermediate grasps are generally used for thin, lightweight objects (Feix et al., 2014b).

In real grasping literature, it has been shown that objects with irregular shapes present the largest variation in grasping approach (Feix et al., 2014b). A similar result was found in this experiment, where Mug, Clamp, Spoon, Marker and Lego showed user grasp choice agreements below 40%. These objects not only have irregular shapes but also present multiple graspable locations such as handles (Mug, Clamp), multiple intuitive graspable possibilities (Spoon, which could be grasped as a regular cylindrical object, or with more precision for eating; Marker, which could be grasped as a regular cylindrical or with more precision for writing) or very small objects (Lego) which are linked to higher variations of grasp types in real environments (Feix et al., 2014b).

Chapter 5 showed that virtual objects were predominantly grasped using Power grasps in 70.97% instances, Precision in 29.02% instances with no Intermediate grasps being collected. A very similar pattern was found in this experiment, where users chose a grasp from the Power grasp category in the majority of instances (68.26%), while choosing Precision grasp types in 27.66% and Intermediate grasp types in only 4.07% instances. This shows that even when additional virtual objects are used, the VR grasping patterns remain consistent. This consistency can also be observed when looking at the most common grasps in this experiment, namely Large Diameter [P1], Precision Disk [PC10], Medium Wrap [P3], Power Sphere [P6] and Small Diameter [P2], which represent the main VR grasps identified in Chapter 5.

Grasp Type / Object Category	Main VR Grasps [85,18%]						Thumb-Finger Variations [7,8%]					Sphere Variations [4,49%]			Other [2,53%]													
	[P1]	[PC10]	[P3]	[P6]	[P2]	[I5]	[PC1]	[PC6]	[PC4]	[PC11]	[PC5]	[PC9]	[PC7]	[PC2]	[I4]	[PC8]	[P5]	[P13]	[P4]	[P14]	[PC12]	[P12]	[P8]	[P7]	[P16]	[I3]	[P10]
Equant	37,27% [432]	1,55% [18]	1,29% [15]	37,70% [437]	0,43% [5]	0,00% [0]	2,85% [33]	0,78% [9]	1,12% [13]	1,81% [21]	0,60% [7]	5,26% [61]	3,02% [35]	2,93% [34]	0,69% [8]	1,90% [22]	0,35% [4]	0,00% [0]	0,09% [1]	0,00% [0]	0,00% [0]	0,00% [0]	0,26% [3]	0,09% [1]	0,00% [0]	0,00% [0]	0,00% [0]
Prolate	0,85% [10]	1,01% [12]	48,52% [574]	0,76% [9]	15,98% [189]	12,76% [151]	3,72% [44]	5,92% [70]	2,70% [32]	1,61% [19]	2,28% [27]	0,08% [1]	0,34% [4]	0,76% [9]	1,44% [17]	0,08% [1]	0,08% [1]	0,76% [9]	0,00% [0]	0,08% [1]	0,00% [0]	0,00% [0]	0,00% [0]	0,08% [1]	0,00% [0]	0,17% [2]	0,00% [0]
Oblate	68,41% [812]	14,49% [172]	3,12% [37]	1,52% [18]	1,26% [15]	0,00% [0]	2,11% [25]	0,67% [8]	0,59% [7]	1,01% [12]	0,59% [7]	0,67% [8]	1,52% [18]	2,02% [24]	0,42% [5]	0,67% [8]	0,42% [5]	0,00% [0]	0,25% [3]	0,17% [2]	0,00% [0]	0,00% [0]	0,00% [0]	0,00% [0]	0,08% [1]	0,00% [0]	0,00% [0]
Bladed	48,32% [574]	40,74% [484]	0,67% [8]	2,53% [30]	1,18% [14]	0,17% [2]	1,18% [14]	1,01% [12]	0,25% [3]	0,17% [2]	0,25% [3]	0,34% [4]	1,09% [13]	0,08% [1]	0,59% [7]	0,42% [5]	0,17% [2]	0,17% [2]	0,00% [0]	0,00% [0]	0,25% [3]	0,25% [3]	0,00% [0]	0,00% [0]	0,08% [1]	0,00% [0]	0,08% [1]

Figure 6.16: VR Taxonomy of Grasp Types. Grasps are categorised by frequency, showing percentage and number of instances for each object category: Equant, Prolate, Oblate and Bladed.

6.8.4 Taxonomy of Grasp Types

The Real Grasp Taxonomy by Feix et al. (Feix et al., 2009, 2014a, 2014b) is structured based on findings from the study of human grasping and by grouping grasp types into Power, Precision, and Intermediate, as well as Thumb Abducted and Thumb Adducted, which are known as taxonomy dimensions, as discussed in Chapter 4. To define the dimensions of the first VR grasping taxonomy, the 27 grasp types used by participants in this experiment were grouped in meaningful categories by frequency of usage in this experiment. The first subcategory is Main VR grasps, which represent the six most used grasps and account for 85.15% of the data (N = 4018)

Results from this study showed that high variations in grasp choice were due to participants using a different number of fingers to perform similar grasps with the same objects. Therefore, these variations were grouped under Thumb-Finger Variations category, which accounted for 7.8% (N = 368) of the total dataset and contains Precision grasp types where the number of fingers used when performing a grasp varies: Thumb-Index Finger [PC1], Thumb 2-Finger [PC4], Thumb 3-Finger [PC5], Thumb 4-Finger [PC6] and Tip Pinch [PC11].

The next category by frequency of use is Sphere Variations which accounted for 4.49% (N = 212) of the total dataset. This category contains Precision grasp types where the hand preshapes in a way similar to grasping a spherical object: Precision Sphere [PC9], Tripod [PC7] and Inferior Pincer [PC2]. The remaining grasp types were grouped in category Other, which was used in only 2.53% instances (N = 119): Lateral Tripod [I4], Quadpod [PC8], Power Disk [P5], Light Tool [P13], Ring [P4], Adducted Thumb [P14], Parallel Extension [PC12], Palmar [P12], Sphere 3-Finger [P8], Sphere 4-Finger [P7], Fixed Hook [P15], Lateral Pinch [I3] and Extension Type [P10]. For a description of these individual grasps please see Figure 6.4.

Figure 6.16 presents the first VR grasping taxonomy which shows correlations between object category and grasp type with percentages and number of instances for each object category (Equant, Prolate, Oblate and Bladed).

6.8.4.1 Analysis - Taxonomy of Grasp Types

Following the user elicitation and labelling methodologies presented in Chapter 4 the first VR Taxonomy of Grasp Types is proposed. The taxonomy was structured following methodologies for defining taxonomy dimensions in real grasping literature (Feix et al., 2009). The six main VR grasps account for more than 85% of the dataset, while a total of 13 different grasp types account for 82.80% of the data in the most complete grasping taxonomy for real objects to date (Feix et al., 2009), showing an overall lower variability in grasping objects of different shapes in VR. This lower variability in grasp approach was also found in grasping virtual objects in immersive technology grasping literature (Al-Kalbani et al., 2016a) and in gesture elicitation studies (Billinghurst et al., 2014), where subjects used a small variety of hand poses across tasks.

Studies looking at grasping patterns in real environments showed that the Medium Wrap [P3] grasp type from the Power grasp category is the most common grasp used when manipulating real objects (Feix et al., 2014b), yet, the VR taxonomy presented in this study showed the Large Diameter [P1] to be the most common grasp type across object categories being used for $\approx 40\%$ of the dataset. The main difference between these two grasps is that Large Diameter [P1] presents a larger hand opening (GAp) than Medium Wrap [P3]. This shows that even though the virtual objects used in this study were of different shapes and sizes subjects did not focus on performing a grasp around the boundaries of the virtual object, which might have been influenced by the lack of haptic feedback as discussed in Chapter 5.

When looking at how grasp type pattern changes between object shape categories, unique patterns were identified for different shapes. Equant objects were predominantly grasped using a Power Sphere [P6] and a Large Diameter [P1] from the Main VR grasps, followed by Sphere variations. This shows that subjects adjusted their grasp to object shape as Equant objects are variations of cuboid and spherical objects. A similar pattern was found for Prolate objects, which are long, tubular bodies and were predominantly grasped using Medium Wrap [P3] and Small Diameter [P2] which represent grasps that differ in terms of hand opening, with both of them being predominantly used to create a stable grasp for heavy objects with a graspable size of 5 to 45 mm in real grasping (Feix et al., 2014b). While the weight of the objects was not an influencer in this case, the tubular shape and the object sizes of the graspable locations (less than 36 mm for every object in this category) might have influenced the grasping pattern for this object category, showing a link between virtual object shape and grasp type, even when hand occlusion and lack of haptic feedback introduce errors in object size estimation (Murcia-López & Steed, 2018).

Subjects used a higher number of finger variations grasps for Prolate category compared to other categories, which might be linked to the lack of haptic feedback (such as weight) associated with VR interactions, that allows users to grasp objects in a comfortable manner instead of prioritising stability (e.g. a hammer from Prolate category can be grasped using a pinch grasp or other finger variations in VR, which would not be possible in real environments). A lack of awareness in the number of fingers involved in grasping was observed for all virtual objects used in this experiment. Variations of the same grasp, but using a different number of digits to perform precision grasps has been used instinctively as shown in the

VR Taxonomy ([PC1], [PC6], [PC4], [PC1] and [PC5]) for the same object. This finding is in connection with prior user elicitation studies defining mid-air gesture interactions for augmented reality (Piumsomboon et al., 2013) and wall display interactions (Wittorf & Jakobsen, 2016) where users did not show awareness of the number of digits they were using while interacting. This contrasts with grasping and manipulating real objects, where the number of digits involved is influenced by the size of the object (Bullock, Feix, & Dollar, 2015), increasing with size and mass (Cesari & Newell, 1999, 2000).

Oblate objects were predominantly grasped using Large Diameter [P1] from the Power category, followed by Precision Disk [PC10] from Precision category. A similar pattern was found for Bladed objects, however the distribution of Large Diameter [P1] and Precision Disk [PC10] is more balanced for Bladed objects. While a high variability in grasp types of the same category was expected, a prevalence for grasp types from different categories, which are fundamentally different from each other introduces the question of whether the virtual objects that propose multiple grasping possibilities (such as Clamp in Oblate and Scissors in Bladed) as presented in Chapter 5, represent in fact outliers that skew the results of the taxonomy for these categories. A post-analysis revealed that more than 98% of the Precision Disk [PC10] grasp instances found in Oblate category were for grasping the Clamp object while more than 90% of the Precision Disk [PC10] grasp instances found in Bladed were for Scissors and Sponge. This change in grasping pattern was also identified in Chapter 5 for objects with irregular shapes and together with results of this study suggest that additional categorisation methods that take these into consideration might provide a more detailed overview of how object characteristics influence grasping patterns in VR.

6.9 Discussion and Conclusions

Results in this chapter showed that virtual object shape influences grasp patterns in VR in terms of grasp labels. Unsurprisingly, GAp did not show significant differences between object shape categories, however it did show to vary for individual objects, suggesting a correlation between the size of the grasped location and GAp when grasping in VR, which is consistent with real grasping literature (Feix et al., 2014b). As found in Chapter 5, subjects grasped virtual objects smaller and larger than object size, with the lack of haptic feedback introducing errors in object size estimation. Objects within the same category showed similar patterns in how GAp relates to object dimensions (A, B and C), with the only exception being the Equant category, where each individual object showed a unique pattern. This might be due to the fact that each object in this category proposes different grasping challenges that might be unrelated to the category they belong to such as the Mug proposing multiple graspable locations (body, top and handle) and Lego being smaller than other objects, which in Chapter 5 has showed to propose a unique pattern in terms of GAp, being grasped larger than all object dimensions. However, to fully understand how GAp relates to object characteristics, the correlation between GAp and grasped location needs to be analysed in more detail. Moreover, to fully understand how grasping patterns change for objects that propose multiple graspable locations, other categorisation methods are explored in Chapter 8.

While a clear pattern in GAp was not identified for Zingg's object categories (Zingg, 1935), a strong correlation between GAp and grasp types can be observed in this work. Equant, Oblate and Bladed objects were predominantly grasped with a $GAp \approx$ 50 - 60 mm and a grasp type Large Diameter [P1] which in real environments is linked to a grasp size of 70 mm, while Prolate objects were grasped with

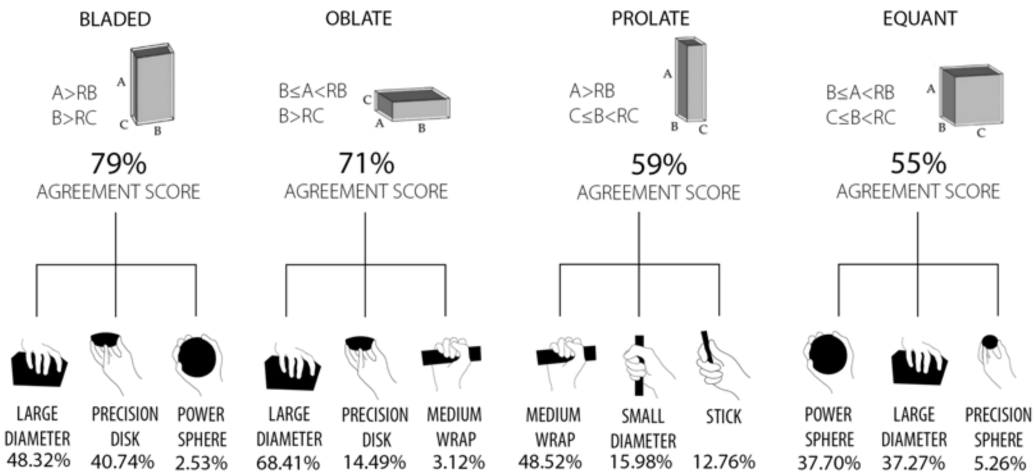


Figure 6.17: Grasping recommendations for virtual objects presenting a grasp choice decision tree based on the most prevalent grasps per object category. Agreement score shows the agreement between participants in choosing a grasp type for each object category.

a $GAp \approx 37$ mm and a grasp type Medium Wrap [P3] which in real environments is linked to a grasp size of 45 mm. Moreover, objects that showed a lower user grasp choice agreement also showed higher variability in GAp, showing that the correlation between positions of the fingers (grasp types) change together with the hand opening (aperture) as known from real grasping. Therefore, the hypothesis H_1**: Virtual object shape does have an effect on grasping patterns in VR** is accepted as differences in grasping labels were found between object shape categories. When comparing grasp labels for different translate tasks, no significant differences were found, therefore failing to reject the null hypothesis H_2**: Translate tasks do not have an effect on grasping patterns in VR**.

This chapter presented the first VR Taxonomy of Grasp Types based on object shape, designed based on existing methodologies for collecting and structuring data in a taxonomy, detailed in Chapter 4. Taxonomies are often summarised in decision trees to provide ways for researchers to quickly identify solutions for interaction problems (Klompmaker et al., 2013), therefore, Figure 6.17 presents a grasp choice decision tree based on the most prevalent grasps per object category. Following Zingg's methodology for object categorisation, objects can be classified based on the relative length of their axes into the four subgroups presented in this chapter. Over 75% of the interactions in each category can be modelled by the three most prevalent grasps, therefore, based on object shape, VR designers could model the most suitable interaction grasps by using this categorisation framework.

7 | Thermal Visual Cues

This work was published in the proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI2020) as "Too hot to handle: An evaluation of the effect of thermal visual representation on user grasping interaction in virtual reality" (Blaga et al., 2020).

7.1 Introduction

One important facilitator of immersion in VR is fidelity, the degree of accuracy with which a system recreates real-world experiences (Bowman, McMahan, & Ragan, 2012; Nabiyouni, Saktheeswaran, Bowman, & Karanth, 2015; Witmer & Singer, 1998), with high levels of interaction fidelity being preferred for virtual object manipulation (Rogers, Funke, Frommel, Stamm, & Weber, 2019).

Multi-sensory feedback has proved to generate high levels of fidelity in VR (Dinh, Walker, Hodges, Song, & Kobayashi, 1999; Fröhlich & Wachsmuth, 2013; Harley et al., 2018), with an increased number of works using haptic feedback devices to stimulate other sensory channels (Achibet, Girard, Talvas, Marchal, & Lećuyer, 2015; Benko, Holz, Sinclair, & Ofek, 2016; Lopes, You, Cheng, Marwecki, & Baudisch, 2017). Yet, providing haptic feedback is still a challenge in VR, which influenced an increased focus on the use of visual cues to represent haptic sensations in VR (Araujo et al., 2016; Lederman & Klatzky, 1987).

Chapter 5 and 6 showed differences in grasping patterns in VR, however, findings also showed that subjects occasionally mimic real grasping behaviour when interacting in VR, specifically when wrapping their hand around the virtual objects. To explore if and how an element of fidelity, specifically thermal representation, influences grasping approach in VR, this chapter presents the first study to evaluate thermal visual cues for grasping virtual objects. Across 50 participants, three thermal cues and four environmental conditions, grasp instances are analysed in terms of grasp aperture (GAp), grasp location and grasp labels.

The chapter is structured as follows: Section 7.2 presents a literature review of existing work on virtual environment representation through visual cues; Section 7.3 presents the experiment design with a detailed methodology being presented for categorising objects, tasks, hand representation and environment; Section 7.4 presents the protocol for the user elicitation study presented in more detail in Chapter 4; Section 7.4.1 presents the metrics used for analysing grasping patterns; Section 7.5 presents the proposed hypotheses; Section 7.6 presents the methodology used for data analysis; Section 7.7 presents results which report on grasping metrics and how visual thermal cues influence grasping approach and Section 7.8 presents discussion and conclusions.

7.2 Background

One important facilitator of immersion in VR is fidelity, the degree of accuracy with which a system recreates real-world experiences (Bowman et al., 2012; Nabiyouni et al., 2015; Witmer & Singer, 1998), with high levels of interaction fidelity being preferred for virtual object manipulation (Rogers et al., 2019).

Multi-sensory feedback has proved to generate high levels of fidelity in VR (Dinh et al., 1999; Fröhlich & Wachsmuth, 2013; Harley et al., 2018), with an increased number of works using haptic feedback devices to stimulate other sensory channels (Achibet et al., 2015; Benko et al., 2016; Lopes et al., 2017). However, providing haptic feedback generally requires complex hardware, while still being limited (Choi et al., 2018; J. Lee et al., 2019).

As an alternative, researchers rely on the concept of the kinesthetic visual capture (Somers & McNally, 2010), the dominance of vision over proprioception, to efficiently deceive user experience through visual cues in the environment (Pillai, Ismail, & Charles, 2017). Rietzler et al. (Rietzler, Geiselhart, Gugenheimer, & Rukzio, 2018) used visual cues to induce the haptic sensation of weight. They used perceivable tracking offsets of the virtual hand, nudging the user to lift the arm higher to perceive some form of additional exertion. To create the illusion of resistance of wind in VR, Pusch et al. (Pusch, Martin, & Coquillart, 2008, 2009) used visual hand displacements. Their results show that the majority of participants felt a force that was pushing their hands. Rosas et al. (Rosas, Wichmann, & Wagemans, 2007) investigated how different types of visual cues for textures change depth perception. They found that textures with a pseudo-random distribution of circles provide the highest reliability in discriminating the distance of objects in motion. Biocca et al. (Biocca, Kim, & Choi, 2001) investigated sensory illusions in a virtual environment, and identified that when manipulating the visual analogue for a physical force, a virtual spring, users reported haptic sensations of *"physical resistance"*, even though the interface included only visual representations. Vigier et al. (Vigier, Moreau, & Siret, 2015) studied the role of visual cues (sky aspect, shadows, sun location and light effects) on climate perception (season, daytime and temperature) in virtual urban environments. Their results prove the feasibility of suggesting complex climatic perceptions and thermal feelings using just visual representations.

When looking at visual representations of the object in VR, several properties can be considered such as weight, texture, however one of the prominent areas currently explored in VR is thermal representation (Z. Chen, Peng, Peiris, & Minamizawa, 2017), which has shown to contribute to immersive user experience (Ranasinghe, Jain, Karwita, Tolley, & Do, 2017) and have been used for a wide range of applications such as training environments (Shaw et al., 2019) and simulations (Ranasinghe et al., 2018). However, visual cues for thermal feedback and their influence on grasping patterns in VR has not been explored before. To address this, this chapter aims to explore the influence of visual cues for thermal haptic feedback on grasping interaction in VR.

7.3 Experiment Design

To analyse the effect of visual cues for temperature on grasping interaction patterns, an elicitation study was designed following the methodology detailed in Chapter 4. This section provides a detailed overview of the apparatus, environment, conditions and 3D hand representation as detailed in Table 4.3.

7.3.1 Apparatus

As in Chapter 5 and 6, a custom experimental framework was built using the Oculus Rift DK2 VR Head Mounted Display (HMD) and the Leap Motion tracking device as detailed in Chapter 4. The Leap Motion was mounted on the front of the HMD facing the participants' hands, to facilitate hand interaction. A web camera was used for recording participants' hand during interaction. The system was developed using C#, Unity 2018.2 and Leap Motion 4.0 SDK.

7.3.2 Environment

The experiment was conducted in a controlled environment under laboratory conditions at location N 52° 28.35′ E 53° 5.87′ in July 2019. The average outdoor temperature was 22.4°C (SD = 3.8)[1] and the indoor controlled environment tem-

[1]Weather data collected from: https://www.accuweather.com/en/gb/birmingham/b5-5/july-weather/326966 *(last accessed September 2019)*

Figure 7.1: Virtual environment showing a virtual desk and a virtual window which changed views in between conditions.

perature was constant 20°C to minimise the potential inference of environmental and weather conditions in the results of the presented study. The test room was lit by a 2700k (*warm white*) fluorescent with no external light source.

Comparable to the environments in Chapter 5 and 6, the virtual environment showed a virtual desk with its surface aligned to a seating position, as in Figure 7.1. A virtual mug was placed at the centre of the table, changing its texture and content as presented in the next subsection and in Figures 7.4(a)-7.4(d). Additionally, the scenario showed a window, which changed views in between conditions.

7.3.3 Conditions

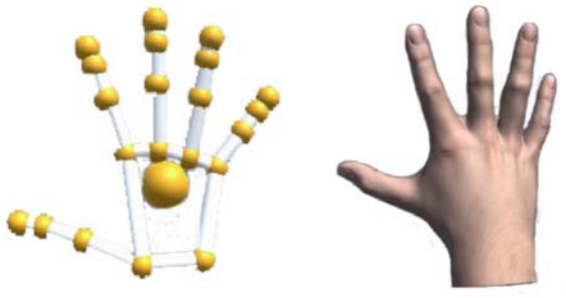
7.3.3.1 Hand Representation

The use of avatars and avatar representation in VR has received significant attention from the research community exploring how it effects the sense of body ownership (Nowak & Biocca, 2003; Serino et al., 2013) and agency (Kilteni, Groten, & Slater, 2012) as they allow users to locate their own body pose within the virtual environment. Furthermore, additional visual aspects such as human-likeness (Lin & Jörg, 2016; Lugrin, Latt, & Latoschik, 2015), gender (Schwind et al., 2017) or transparency (Buchmann, Nilsen, & Billinghurst, 2005; Knierim, Schwind, Feit, Nieuwenhuizen, & Henze, 2018) of the virtual representation have shown to influence ownership illusion as well as user performance.

The human hand is a powerful physical tool through which people interact with the surrounding world (Flanagan & Johansson, 2001). It has been identified that perceived naturalness of virtual hands can have a significant effect on perceived user presence (Schwind et al., 2017) as well as own-body perception and immersion (Lin & Jörg, 2016). Therefore, for hand representation two hand models were used: *abstract* hand model from the Leap Motion SDK used in Chapter 5 and 6, which was represented as a set of cylinders and spheres representing bones and joints respectively (please refer to Figure 7.2(a)) and *human* hand model, an androgynous hand model chosen following Schwind et al. (Schwind et al., 2017) recommendation for avoiding noticeable gender characteristics in human hands (please refer to Figure 7.2(b)).

7.3.3.2 Thermal Representations

A virtual mug was chosen from the object set presented in Chapter 4, due to its familiarity in everyday tasks, as well as the grasp variations it proposes. Won



(a) Abstract (b) Human

Figure 7.2: Hand representation showing a) abstract hand model and b) human hand model

et al. (Won & Westland, 2017) showed that the colour red is associated with hot concepts, blue with cold and green with reliable and safe. For the mug object used in this experiment, the colour yellow was used as it is the only one of the primary colours that was not associated to any connotation that may influence the study results.

(a) Cold (b) Hot (c) Empty

Figure 7.3: Thermal representations for the virtual mug. 7.3(a) shows the cold condition with ice cubes and a clear liquid, 7.3(b) shows the hot condition with coffee steam coming out of the top and 7.3(c) shows the empty condition with no content.

- **Cold:** The mug content in this condition was a set of ice cubes inside a clear blue liquid (as in Figure 7.3(a)).
- **Hot:** The mug content was rendered as a brown liquid simulating coffee

with steam coming from the top of the mug (as in Figure 7.3(b)).

- **Empty:** An empty mug with no content inside (as in Figure 7.3(c)).

The colours of the liquid (blue for cold and brown for hot) were chosen based on previous literature showing that surfaces whose dominant frequencies are towards the blue end of the spectrum are perceived as cold, and those towards the red end of the spectrum are perceived as hot (C. A. Bennett & Rey, 1972).

7.3.3.3 Environmental Cues

The environment surrounding objects has shown to influence interaction choice (Wimmer, 2011), therefore, to evaluate this in the context of the study, different contextual representations were explored to support the thermal cues above.

- **Basic:** Presents a simple yellow mug, in all three temperature conditions. As in Figure 7.4(a) the only visual difference between the thermal conditions is the rendered content inside the mug.

- **Content Label:** Presents a mug with a label attached to it, informing about the contents inside as in Figure 7.4(b).

- **Glass:** Presents a see-through mug in a transparent texture as in Figure 7.4(c). This allows the user to see the content of the mug through the mug itself, mimicking a glass texture.

- **Context Objects:** Presents the mug used in the basic condition accompanied by other contextual objects to support the thermal representation as in Figure 7.4(d). These accompanying objects were presented behind the mug, 10 cm away from its original position in both Z and X axes. The accompanying items were a coffee espresso machine for the hot condition, ice bucket for the cold condition and an empty bucket for the empty condition.

The view from the window also changed depending on the thermal cues, displaying a snowy landscape for the hot condition, a beach for the cold condition and a forest for the empty condition.

(a) Basic (b) Label (c) Glass (d) Context Objects

Figure 7.4: Conditions under study, with 7.4(a), 7.4(b), 7.4(c) and 7.4(d) showcasing the environmental cues for each visual thermal condition: hot, cold and empty.

7.3.4 Task

As with Chapter 5, a simple translation task on one direction is introduced. Participants were instructed to pick and move the virtual mug from the origin location to a marker location situated on the left of the original object and displayed as a 3D semi-transparent virtual mug in a different colour, as in Figure 7.5.

7.3.5 Participants

A total of 50 participants (21 females, 29 males) from a population of university students and staff members volunteered to take part in this study. Participants ranged in age from 18 to 50 (M = 25.5, SD = 14.57). All participants were right-

Figure 7.5: Interaction environment displaying the virtual mug (yellow virtual mug) and the marker position (semi-transparent green virtual mug).

handed, to ensure they interacted with the mug under the same conditions (i.e. the handle in the same orientation with respect to their dominant hand).

All participants performed the experiment tasks under both hand conditions (abstract and human). Comparable to Chapter 5 and 6, participants completed a standardised consent form and were not compensated. Visual acuity of participants was measured using a Snellen chart, each participant was also required to pass an Ishihara test to check for colour blindness. Participants with colour blindness and/or non corrected visual acuity of < 0.80 (where 20/20 is 1.0) were not included in this study. Participants were asked to self-assess their level of experience with VR systems, with 16 participants reporting to have an average level of experience, 31 reported being novice to the technology and 3 self-labelled themselves as experts. Participants did not have any previous experience with hand tracking sensors.

7.4 Protocol

7.4.0.1 Pre-test

Prior to the study, participants were given a written consent form, where the test protocol and main aim of the study were described. Additionally, participants completed a pre-test questionnaire enquiring about their background level of experience with VR systems and hand recognition sensors.

7.4.0.2 Training

Participants were trained to pick and move a neutral object (a cube) from its original position to a target position as in Chapter 5 and 6, to familiarise themselves with the VR environment and hand interaction space. Thermal cues were not included at this stage of the study. Participants spent 7-10 minutes training with the system until they felt comfortable with the task and the apparatus.

7.4.0.3 Test

Once participants were comfortable with the interaction space and the overall VR environment, they were presented with the main experimental task. For each task, participants were instructed to be seated in a neutral position with their hands at the side, while the virtual mug under experiment appeared in the starting position. Each participant completed 24 tasks (2 hand representations $\times$ 3 temperature cues $\times$ 4 environmental conditions) with a total of 1200 grasps being collected during the experiment (50 $\times$ 24). The order of the hand representation conditions were counterbalanced; half participants started with human hand (Figure 7.2(b)) and the other half with the abstract hand (Figure 7.2(a)). Thermal and environmental conditions were then presented in randomised controlled order. Participants were asked to pick and move the virtual mug the way it felt most intuitive to them and instruct the test coordinator when they were happy with their grasp.

7.4.0.4 Post-test

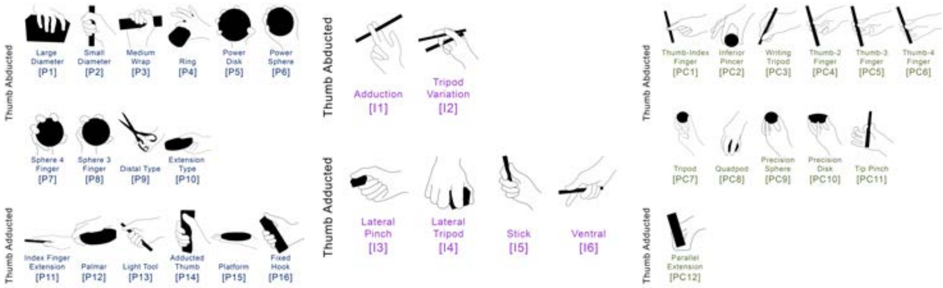
As no motion sicknes or cognitive overload was reported in Chapter 6 and because the same environment configuration and simple translation tasks are used in this experiment, MSAQ and NASA-TLX are not assessed in this experiment. Instead, participants were asked to complete the Igroup Presence Questionnaire (IPQ) and a set of tailored questions asking about their experience during interaction with the virtual object in different conditions.

7.4.1 Metrics

Grasp Aperture (GAp**):** GAp is a metric used to measure how accurately users estimate the size of a virtual object and has been described in more detail in Chapter 4 and calculated in the experiments presented in Chapter 5 and 6.

Grasp Labels: As with Chapters 5 and 6, the grasp categories are Power, Intermediate and Precision. Power grasps are linked to stability and security (Figure 7.6 a). Intermediate grasps present elements of Power and Precision roughly in the same proportion, enabling a finer representation of grasp types (Bullock et al., 2013) (Figure 7.6 b). Precision grasps present grasps where the object is commonly held between the fingertips (Figure 7.6 c).

Grasp Location: An object can be manipulated in different ways. For each way it is manipulated, there might be different proportions of the object relevant for the actual grasp. Therefore, Feix et. al. (Feix et al., 2014b) introduced the concept of grasped location, which they define as the local part of the object specific to



(a) Power Grasps (b) Intermediate Grasps (c) Precision Grasps

Figure 7.6: Grasp types from the real grasp taxonomy (Feix et al., 2009)

the grasp instance. Therefore, for the virtual object used in this study (mug), three grasp locations can be defined as shown in Figure 7.7: Body/Side, Top and Handle.

Figure 7.7: Grasp locations described by (Feix et al., 2014b) for the virtual mug used in this study.

Presence Questionnaire: The IPQ is a scale for measuring the sense of presence experienced and was used to understand how hand representation and visual cues for thermal haptic feedback influence presence while grasping in VR. When compared to other presence questionnaires, IPQ has shown to provide the highest reliability (Schwind, Knierim, Haas, & Henze, 2019). The questionnaire is structured in 4 sub-scales: General Presence (PRES), Spatial Presence (SP), Involvement (INV) and Experience Realism (REAL), with 14 items in total, rated in

7-point scale (1-no feeling of presence, 7-strong feeling of presence). The scores for each sub-scale as well as the overall score are calculated by averaging their 7-point scores.

Post-Test Questionnaire: The post-test questionnaire consisted of tailored questions as follows:

- Which of the below mostly influenced the way you interacted with the mug? [Visual cues for temperature, Warning labels, Mug material, Other environmental cues, None]
- Please rate the perceived usefulness of the following visual cues for choosing a grasp: [Steam] (1- Not useful, 5-Very useful)
- Please rate the perceived usefulness of the following visual cues for choosing a grasp: [Ice] (1- Not useful, 5-Very useful)
- Please rate the perceived usefulness of the following visual cues for choosing a grasp: [Warning labels on the mug] (1- Not useful, 5-Very useful)
- Please rate the perceived usefulness of the following visual cues for choosing a grasp: [Environmental cues] (1- Not useful, 5-Very useful)
- Do you think you interacted differently with the mug that had content (coffee and water) compared to the empty mug?
- Do you think temperature representations influenced the location where you grasped the mug (handle or the body of the mug)?

7.5 Hypotheses

While assumptions can be made from the literature, that the influence of visual cues in the virtual environment can have an effect on user perception, no literature exists to suggest the influence it has on grasping metrics, therefore the following null hypothesis is proposed:

H_1**:** The thermal visual cues of the object have no effect on the grasp metrics.

To understand if grasping patterns change for different hand representations, which is currently unknown in VR, the following null hypothesis is proposed:

H_2**:** The visual representation of the hand has no effect on the grasp metrics.

7.6 Data Analysis

The Shapiro-Wilk (Shapiro & Wilk, 1965) normality test found the data to be not normally distributed. Data collected for GAp was non-parametric and not normally distributed, therefore statistical significance between the dependent groups (Empty, Hot and Cold) where the variable of interest is continuous (GAp in mm) was tested using the Friedman test (Friedman, 1940) with an alpha of 5%. For testing statistical significance between grasp patterns, where the dependent variable (Grasp Location and Grasp Category) is nominal categorical (Body, Top and Handle; Power, Intermediate and Precision), contingency tables were created and analysed for significance using a Chi-Squared Test of Independence with 95% Confidence Intervals, therefore, a p-value of less than .05 will indicate statistical significance. Cramer's V calculation for effect sizes was applied after verifying its assumptions that the variables under analysis are categorical. Results were interpreted following existing guidelines based on degrees of freedom.

7.7 Results

7.7.1 Grasp Aperture (GAp)

Grasp Aperture (GAp) was analysed per temperature representations and scenario conditions to understand the accuracy with which users estimated the size of the grasped object.

7.7.1.1 Hand Representation

To understand if hand representation influenced GAp when grasping virtual objects, differences in GAp between abstract and human hand were statistically analysed with 95% Confidence Intervals. No statistical significance was found in GAp for abstract (M = 42.90, SD = 27.94) and human hand representation (M = 43.99, SD = 30.77): χ^2 *(1, N = 1200) = 88579.01, p = .712*.

7.7.1.2 Environmental Conditions

Statistical significance was tested between temperature representations (Empty, Hot and Cold) for every scenario condition (Basic, Content Label, Glass and Context Objects) and hand representation (Abstract and Human) with 95% Confidence Intervals. Pair-wise Effect Sizes (ES) are reported for each condition.

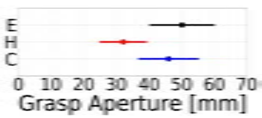
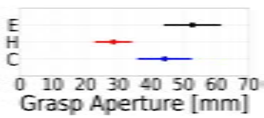
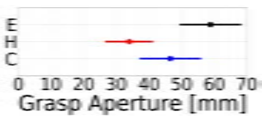
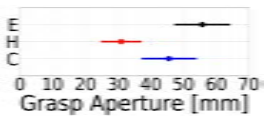
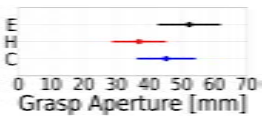
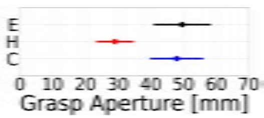
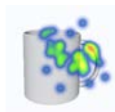
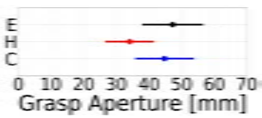
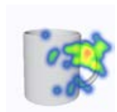
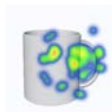
- Basic: Statistically significant differences were found between temperature representations for abstract hand with hot condition (M = 32.06, SD = 24.60), cold condition (M = 50.11, SD = 31.52) and empty condition (M = 45.97, SD = 34.07): χ^2 **(2, N = 150) = 8.68, p = 0.01***, with pair-wise ES: Empty vs Hot = 0.67; Empty vs Cold = 0.13 and Hot vs Cold = 0.50. Statistically significant differences were also found for human hand with hot condition (M = 28.54, SD = 18.97), cold condition (M = 44.21, SD =

28.20) and empty condition (M = 52.71, SD = 29.64): χ^2 **(2, N = 150) = 26.28, p < 0.01***, with pair-wise ES: Empty vs Hot = 0.95; Empty vs Cold = 0.27 and Hot vs Cold = 0.54.

- Content Label: Statistically significant differences were found between temperature representations for abstract hand: hot condition (M = 33.90, SD = 24.93), cold condition (M = 46.52, SD = 32.14) and empty condition (M = 58.79, SD = 32.31): χ^2 **(2, N = 150) = 13.0, p = 0.02***, with pair-wise ES: Empty vs Hot = 0.86; Empty vs Cold = 0.38 and Hot vs Cold = 0.44. Statistically significant differences were also found for human hand: hot condition (M = 30.94 mm, SD = 20.87), cold condition (M = 45.55, SD = 28.51) and empty condition (M = 55.72, SD = 29.07): χ^2 **(2, N = 150) = 22.84, p < 0.01***, with pair-wise ES: Empty vs Hot = 0.96; Empty vs Cold = 0.35 and Hot vs Cold = 0.58.

- Glass: Statistically significant differences were found between temperature representations for human hand: hot condition (M = 29.04, SD = 19.07), cold condition (M = 47.92, SD = 27.61) and empty condition (M = 49.57, SD = 29.80): χ^2 **(2, N = 150) = 7.8 p = 0.02***, with pair-wise ES: Empty vs Hot = 0.82; Empty vs Cold = 0.05 and Hot vs Cold = 0.79. No significance was found between temperature representations for abstract hand: hot condition (M = 36.96, SD = 28.06), cold condition (M = 46.04, SD = 30.63) and empty condition (M = 52.29, SD = 32.82):χ^2 (2, N = 150) = 1.74, p = 0.06, with pair-wise ES: Empty vs Hot = 0.50; Empty vs Cold = 0.22 and Hot vs Cold = 0.28.

Table 7.1: Grasp aperture (GAp) in mm for every temperature representation (H stands for Hot, C stands for Cold and E stands for Empty) with Standard Error (SE). Heat-maps represent the locations where users grasped in each condition for a correlation between GAp and grasped location.

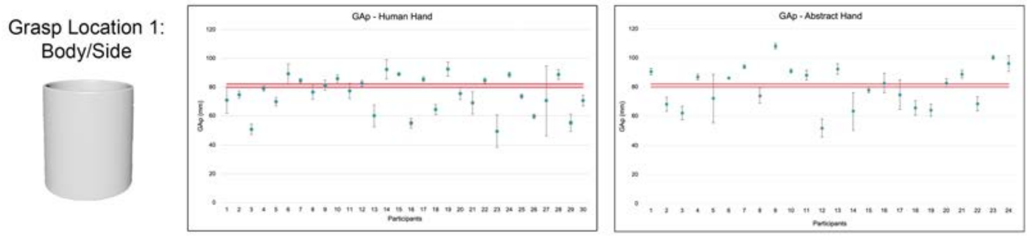
Condition	Abstract		Human	
	GAp	Heat-maps	GAp	Heat-maps
Basic		(a) H  (b) C  (c) E 		(d) H  (e) C  (f) E 
Content Label		(g) H  (h) C  (i) E 		(j) H  (k) C  (l) E 
Glass		(m) H  (n) C  (o) E 		(p) H  (q) C  (r) E 
Context Objects		(s) H  (t) C  (u) E 		(v) H  (w) C  (x) E 

- Context Objects: Statistically significant differences were found between temperature representations for human hand: hot condition (M = 31.56, SD = 22.86), cold condition (M = 43.88, SD = 29.83) and empty condition (M = 55.15, SD = 28.66): χ^2 (2, N = 150) = 11.68 p = 0.03*, with pair-wise ES: Empty vs Hot = 0.90; Empty vs Cold = 0.38 and Hot vs Cold = 0.46. No significance was found between temperature representations for abstract hand: hot condition (M = 34.05, SD = 25.27), cold condition (M = 45.01, SD = 30.50) and empty condition (M = 47.23, SD = 31.77): χ^2 (2, N = 150) = 1.49, p = 0.47, with pair-wise ES: Empty vs Hot = 0.46; Empty vs Cold = 0.08 and Hot vs Cold = 0.38.

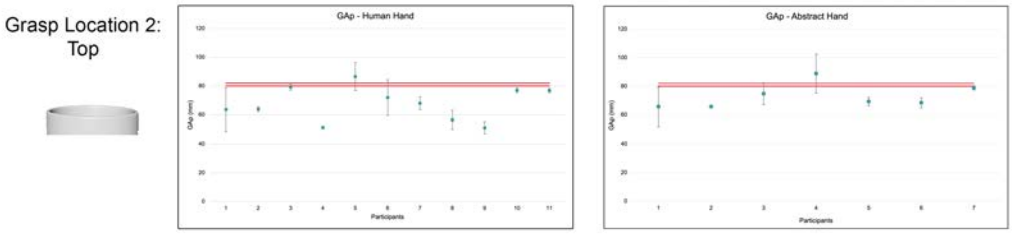
7.7.1.3 Analysis - GAp

The hot condition presented the smallest mean GAp in every environmental condition while the empty condition presented the largest mean GAp in every environmental condition, showing that visual cues for temperature influenced GAp in VR. This is consistent with real grasping behaviour, where humans are more likely to grasp the handle of a mug when it has hot content inside. No statistical differences in GAp were found between hand representations, showing that the realism of the virtual hand did not influence the accuracy with which participants estimated object size when adjusting their hand opening for grasping.

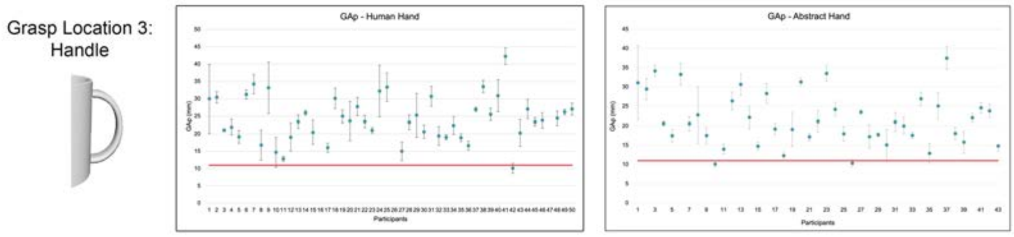
Chapter 6 showed that GAp in VR grasping might be directly influenced by the size of the object at the grasping point, therefore, the grasping point was computed from sensor data as the middle point of the grasp aperture as described in (Al-Kalbani et al., 2016a) and represented in heat-maps to provide an overview of where the mug was grasped for each condition (see Table 7.1). The heat-maps show a prevalence for grasping the handle in hot and cold condition and a prevalence for grasping the body in the empty condition. For example, looking at Table



(a) GAp for Body Location



(b) GAp for Top Location



(c) GAp for Handle Location

Figure 7.8: Overview of GAp of every participant for each grasped location under analysis: Body/Side, Top and Handle. The red lines represent object dimensions (x,y,z).

7.1 Human - Glass condition, in hot condition the concentrated areas are around the handle, in cold condition are predominantly around the handle but also top and body and in empty condition, the concentrated areas are around all three locations of the object. To further analyse the link between object size and GAp, Figure 7.8 shows the GAp of every participant for each grasping instance, structured based on grasped location: Body/Side, Top and Handle. The red lines represent the size of the grasped location (X, Y, Z) in mm. It can be observed that for Body/Side location, there is high variability in GAp, participants grasping both smaller and larger than object dimensions for both hand representations; for Top location, participants generally grasped smaller than object dimensions for both hand representations and for Handle location the majority of users grasped larger than object dimensions for both hand representations.

While results in this work show that there is a link between thermal cues and GAp, which is naturally attributed to the grasped object/location (e.g smaller GAp for hot condition than empty condition due to users grasping the handle more often in the hot condition), GAp patterns showed a high variability in VR, users generally grasping both larger and smaller than object sizes. However, smaller objects seem to be associated with a specific pattern, being grasped larger than object dimensions, which was also found in Chapter 5.

7.7.2 Grasp Location

Grasp location results are reported for each condition based on number of grasp instances for each grasped location: Body/Side, Top and Handle.

7.7.2.1 Hand Representation

To understand if hand representation had an influence on grasp location, differences in grasped location between abstract and human hand were statistically analysed with 95% Confidence Intervals. No statistical significance was found in grasp location for abstract (355 instances for Handle; 198 instances for Body/Side; 47 instances for Top) and human hand representation (384 instances for Handle; 174 instances for Body/Side; 42 instances for Top) :χ^2 *(1, N = 1200) = 2.96, p = .226*.

Environment	Condition	Abstract				Human			
		Handle	Body	Top	Stat	Handle	Body	Top	Stat
Basic	H C E	40 27 23	8 18 22	2 5 5	**Stat = 13.26 p =.01***	45 28 22	8 20 21	1 2 7	**Stat = 27.32 p < .01***
Content Label	H C E	39 29 18	8 19 22	3 2 10	**Stat = 21.95 p < .01***	44 30 20	5 4 21	1 6 9	**Stat = 25.05 p < .01***
Glass	H C E	37 28 22	10 19 24	3 3 4	**Stat = 9.82 p < .01***	45 29 28	4 17 18	1 4 4	**Stat = 16.73 p < .01***
Contextual objects	H C E	38 28 26	9 19 20	3 3 4	Stat = 7.52 p = .111	40 30 23	9 18 20	1 2 7	**Stat = 15.29 p < .01***

Table 7.2: Grasp location results for each environmental condition and thermal representation (H - Hot, C - Cold, E - Empty) presented with statistical results for each hand representation (Abstract and Human).

7.7.2.2 Environmental Conditions

Statistical significance was tested between temperature representations (Empty, Hot and Cold) for every scenario condition (Basic, Content Label, Glass and Context Objects) and hand representation (Abstract and Human) with 95% Confidence Intervals. Effect sizes (ES) are computed using Cramer's V and interpreted using existing guidelines.

- Basic: Statistically significant differences were found between temperature representations for abstract hand where hot showed 40 grasps for Handle, 8 for Body/Side and 2 for Top; cold showed 27 grasps for Handle, 18 for Body/Side and 5 for Top and empty showed 23 grasps for Handle, 22 for Body/Side and 5 for Top: χ^2 **(4, N = 150) = 13.26, p = .01*** with medium ES (Cramer's V = 0.21). Statistically significant differences were also found between temperature representations for human hand where hot showed 45 grasps for Handle, 4 for Body/Side and 1 for Top; cold showed 28 grasps for Handle, 20 for Body/Side and 2 for Top and empty showed 22 grasps for Handle, 21 for Body/Side and 7 for Top: χ^2 **(4, N = 150) = 27.32, p < .01*** with medium ES (Cramer's V = 0.31). The results in terms of grasp location for basic condition can be visualised in Figure 7.9.

- Content Label: Statistically significant differences were found between temperature representations for abstract hand where hot showed 39 grasps for Handle, 8 for Body/Side and 3 for Top; cold showed 29 grasps for Handle, 19 for Body/Side and 2 for Top and empty showed 22 for Body/Side, 18 for Handle and 10 for Top: χ^2 **(4, N = 150) = 21.95, p < .01*** with medium ES (Cramer's V = 0.27). Statistically significant differences were also found between temperature representations for human hand where hot showed 44 grasps for Handle, 5 for Body/Side and 1 for Top; cold showed 30 grasps for Handle, 4 for Body/Side and 6 for Top and empty showed 21 grasps for Body/Side, 20 grasps for Handle and 9 for Top: χ^2 **(4, N = 150) = 25.05, p < .01*** with medium ES (Cramer's V = 0.28). The results in terms of grasp location for content label condition can be visualised in Figure 7.9.

- Glass: Statistically significant differences were found between temperature representations for abstract hand where hot showed 37 grasps for Handle, 10 for Body/Side and 3 for Top; cold showed 28 grasps for Handle, 19 for Body/Side and 3 for Top and empty showed 24 for Body/Side, 22 for Handle and 4 for Top: χ^2 **(4, N = 150) = 9.82, p < .01*** with small ES (Cramer's V = 0.18). Statistically significant differences were also found between temperature representations for human hand where hot showed 45 grasps for Handle, 4 for Body/Side and 1 for Top; cold showed 29 grasps for Handle, 17 for Body/Side and 4 for Top and empty showed 28 grasps for Handle, 18 for Body/Side and 4 for Top: χ^2 **(4, N = 150) = 16.73, p < .01*** with medium ES (Cramer's V = 0.23). The results in terms of grasp location for glass condition can be visualised in Figure 7.9.

- Contextual Objects: Statistically significant differences were found between temperature representations for human hand where hot showed 40 grasps for Handle, 9 for Body/Side and 1 for Top, cold showed 30 grasps for Handle, 18 for Body/Side and 2 for Top and empty showed 23 grasps for Handle, 20 for Body/Side and 7 for Top: χ^2 **(4, N = 150) = 15.29, p < .01*** with medium ES (Cramer's V = 0.22). No significance was found between temperature representations for abstract hand where hot showed 38 grasps for Handle, 9 for Body/Side and 3 for Top; cold showed 28 grasps for Handle, 19 for Body/Side and 3 for Top and empty showed 26 grasps for Handle, 20 for Body/Side and 4 for Top: χ^2 *(4, N = 150) = 7.52, p = .111*. The results in terms of grasp location for contextual objects condition can be visualised in Figure 7.9.

Results above showed that temperature cues influenced the grasped location in VR. To understand if grasped location was also influenced by contextual information, statistical significance was tested between scenario conditions (Basic, Content Label, Glass and Context Objects) in terms of grasp location for both hand representations. No statistical difference was found between scenario conditions in terms of grasped location for abstract hand χ^2 *(4, N = 600) = 2.54, p = .862* and human hand χ^2 *(4, N = 600) = 5.59, p = .469*.

7.7.2.3 Analysis - Grasp Location

Results showed no statistically significant differences in terms of grasp location between hand representations, however, statistically significant differences were found between thermal representations with hot and cold conditions for all environmental conditions except contextual objects condition for abstract hand representation. Results predominantly showed grasps around the handle of the mug for hot condition, while empty condition predominantly showing grasps around the body of the mug (see Figure 7.9). This is consistent with results plotted from sen-

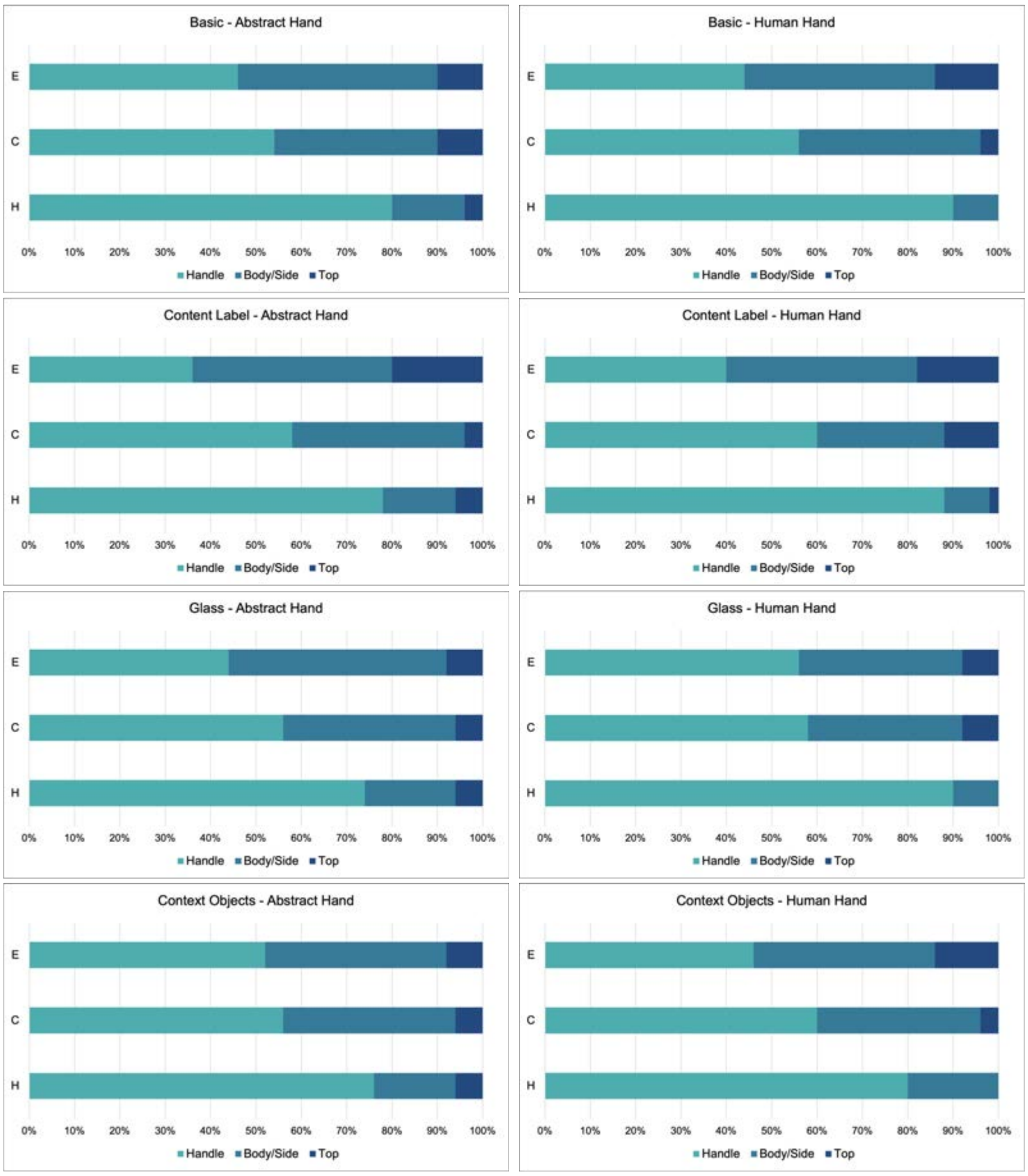


Figure 7.9: Grasp Location (Body/Side, Handle and Top) chosen by participants in this experiment for each contextual environment condition (Basic, Content Label, Glass and Context Objects) and temperature representations (H stands for hot, C stands for Cold and E stands for empty).

sor data presented in the previous section and therefore confirms the assumption made in Chapter 6 that GAp is directly influenced by the size of the grasped location in VR, subjects mimicking real grasping behaviour and adjusting the hand opening for each grasped object: for hot condition where subjects predominantly grasped around the handle, the mean GAp recorded was $\approx$ 30 mm, while for empty condition where subjects grasped around the body, the mean GAp recorded was $\approx$ 50 mm. This finding also shows that participants over and underestimated the size of the grasped object, with the handle having a width of 11 mm and the body a width of 80 mm, which is consistent with findings in Chapter 6.

Additional environmental cues (labels, glass, other objects) did not change the effect of the thermal visual cues, apart from the contextual objects condition with abstract hand, where there were no significant differences between thermal conditions in terms of grasp location, which might have been due to users being distracted by the other objects and not focusing on the thermal visual cues. However, results presented that thermal visual cues showed significant differences in grasp location even when no contextual information was added (basic condition). This is consistent with VR literature showing that subjects use sensory cues in one modality to fill in the missing components of perceptual experience (Biocca et al., 2001), meaning that in this case the visual cues for temperature were enough for users to start behaving as when real haptic thermal feedback was available. This effect of providing haptic sensations without the actual matching haptic stimulus, but instead by inducing those sensations using vision is known as pseudo-haptic feedback and has been highly used in VR literature (Zenner & Krüger, 2017).

Cold and empty conditions showed a similar pattern in grasped location, showing that content inside the mug (suggesting a different weight) did not change the grasping approach, which is consistent with real grasping behaviour where humans generally use the handle when the mug has hot content, and the body/top of the mug for other types of content. While there is a clear link between grasped location and visual thermal conditions in VR, the empty mug (without visual thermal cues) also showed grasps for all locations in all environmental cues. This could imply that more complex virtual objects (that have handles or protruding components) show higher variability in grasping patterns as they do in real environments, which was also discussed in Chapter 6. However, this might also be the result of bias introduced by the thermal conditions explored in this experiment.

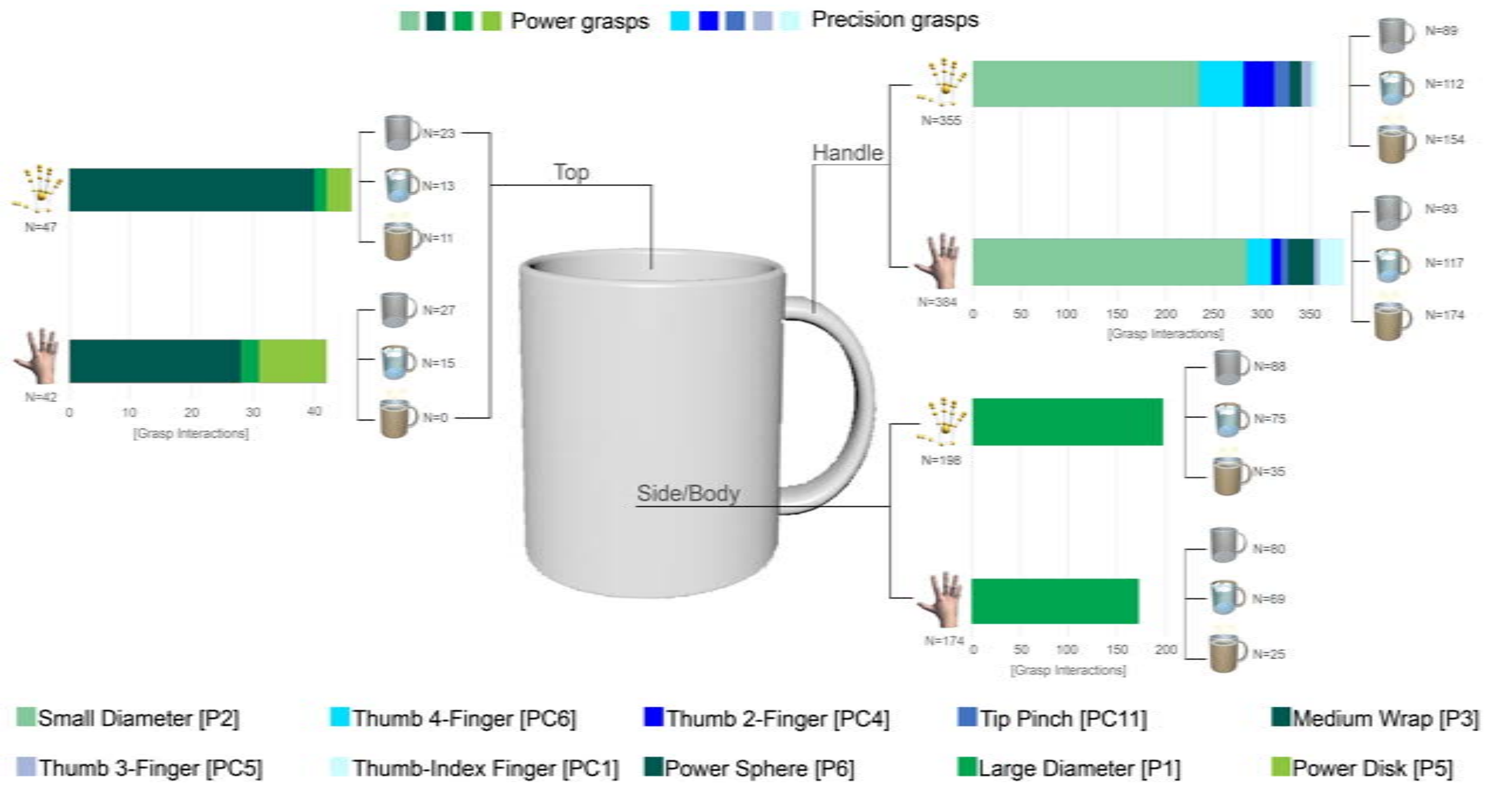


Figure 7.10: Grasp type choice for each grasped location; N represents the number of instances for which that grasp location was chosen, for each temperature condition. Grasp types are categorised in *Power* (variations of green) and *Precision* (variations of blue).

7.7.3 Grasp Labels

A total of 1200 grasps were recorded during the experiments (50 participants x 2 hand representations x 3 thermal cues x 4 environmental conditions) which were labelled following the methodology presented in Chapter 4. No grasps were removed during the labelling. Cohen's Kappa was used to measure inter-rater reliability for labelling the grasps. Raters agreed in 89% of instances (Cohen's Kappa = 0.43) which based on existing guidelines is a moderate agreement, which is often achieved when subjectivity is involved in the process (S. Sun, 2011) and is a common agreement score for classification tasks (Feix et al., 2014b).

Power grasps were the most common grasp types used in this experiment for both abstract (81.83%, N = 491) and human (87.66%, N = 526) hand representations, followed by Precision grasps: abstract (18.16%, N = 109) and human (12.33%, N = 74).

Previous sections showed a link between GAp and grasped location. Chapter 6 showed a link between the shape of an object and the grasping pattern. To understand how every part of the object was grasped, grasp labels are reported for each grasped location: Body/Side, Top and Handle. Body/Side was grasped using only Power grasps for both hand representations: abstract (N = 198) and human (N = 174). A similar pattern was found for Top which was grasped using only Power grasps for both hand representations: abstract (N = 47) and human (N = 42). Handle showed a different pattern being grasped by Power (abstract: 69.29%, N = 246 and human: 80.72%, N = 310) and Precision grasps (abstract: 30.71%, N = 109 and human: 19.38%, N = 74).

As shown in Figure 7.10 Body/Side was predominantly grasped using Large Diameter [P1] for both abstract (100%, N = 198) and human hand (100%, N = 174)

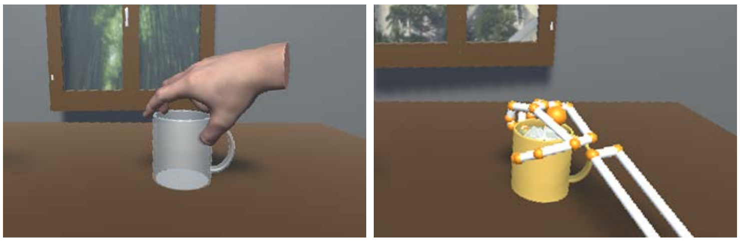
Location	Main Grasp Types		
Grasp Location 1: Body/Side 	 **[P1]** *100%*		
Grasp Location 2: Top 	 **[P6]** *76.40%*	 **[P5]** *16.85%*	 **[P1]** *5.61%*
Grasp Location 3: Handle 	 **[P3]** *70.09%*	 **[PC6]** *9.60%*	 **[PC4]** *5.68%*

Table 7.3: Main Grasp Types chosen in this experiment for each object location (Body/Side, Top and Handle) along with their grasp code (presented in Figure 7.10 and detailed in Chapter 4, colour-coded to outline their grasp category Power grasps in blue and Precision grasps in green.

representations; Top was predominantly grasped using Power Sphere [P6] for both abstract (85.10%, N = 40) and human hand (66.67%, N = 28) representations and Handle was predominantly grasped using Small Diameter [P2] for both abstract (65.91%, N = 234) and human hand (73.95%, N = 284) representations. Table 7.3 shows the main grasp types for each grasp dimension.

7.7.3.1 Analysis - Grasp Labels

Results showed that each grasp location was associated with different grasp labels. This is consistent with results from Chapter 6 showing that the shape of the virtual object influences the grasp type chosen during interaction. In Chapter 6, the Mug object was categorised as Equant and associated with Power Sphere [P6] and Large Diameter [P1]. Results of this study show that the Mug was grasped



(a) Human Hand - Power Disk (b) Abstract Hand - Power Sphere

Figure 7.11: Example of Power Disk [P5] in the human hand condition and Power Sphere [P6] in the abstract hand condition.

using a Power Sphere [P6] when it was grasped by the Top, a Large Diameter [P1] when it was grasped by Body/Side and Small Diameter [P2] when it was grasped by Handle. This shows that the higher variability for irregular objects discussed in Chapter 6 might be due to users grasping on different locations and adjusting their grasp accordingly as in real grasping literature (Feix et al., 2014b).

As shown in Chapter 6, participants did not show an awareness of the number of fingers involved while interacting with the virtual mug. Commonly, users presented different variations of the thumb-finger group presented in the VR Taxonomy of Grasp Types for grasping the handle of the mug. While the prevalence for Power grasps is maintained for grasps recorded in this experiment, differences in grasp types were observed between the human and abstract hand conditions: for instances where the handle was chosen, participants used a Power grasp in 80.72% of the instances for the human hand condition while choosing Power grasps in only 69.29 % of the instances in the abstract hand condition. This result suggests that users intuitively performed a grasp that could normally hold a heavy object more often with the human hand than with the abstract hand, with Power grasps being associated with stability and security when grasping real objects (M. R. Cutkosky, 1989). A similar pattern can be observed for Top, where for both conditions the main grasp types used were Power Sphere [P6] and Power Disk [P5], however with the human hand, the number of instances where Power Disk [P5] was used is higher than with the abstract hand. The difference between these grasps is that Power Sphere [P6] is used for grasping a spherical object where the fingers wrap around the object, while Power Disk [P5] is used for grasping a disk, where the fingertips are the main parts of the fingers that exert forces on the object, to create a stable grasp. In real environments humans generally grasp the top of a glass or mug using a Precision Disk [P5], therefore results in this work show that with the human hand there were more instances where subjects mimicked real grasping behaviour. However, the general trend for both hands was to grasp using a Power Sphere [P6] where the hand penetrates the virtual object as shown in Figure 7.11.

7.7.4 IPQ

The IPQ Presence Questionnaire showed no statistically significant differences between the hand representations under study, with the abstract hand obtaining an average score of (M = 4.54, SD = 1.02) and the human hand (M = 4.75, SD = 1). None of the sub-scales of the questionnaire showed any statistically significant differences. Presence scores by sub-scale are presented in Figure 7.12.

7.7.5 Post-test Questionnaire

Participants were asked to complete a post-test questionnaire to gain a better understanding of their perceptions while interacting with the thermal cues, the environment and hand representations.

Participants reported *ice* and *steam* visual cues from the thermal representations as the strongest visual cue supporting their grasping behaviour. Following these remarks, participants were asked if they felt their grasp location and type was influenced by the different thermal representations or the environment. 37 par-

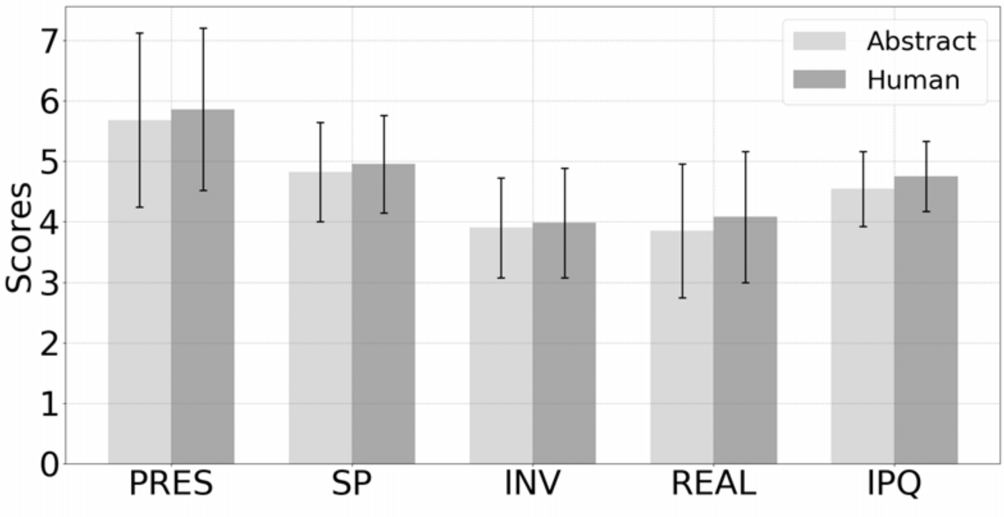


Figure 7.12: Scores for IPQ sub-scales and overall IPQ score for *abstract* and *human* hand conditions; a score equal to 7 represents the highest feeling of presence while 1 represents the lowest.

ticipants (*abstract* hand) and 36 participants (*human* hand) reported that grasp location was influenced by mug content. Additionally, 29 participants (*abstract* hand) and 37 (*human* hand) reported that the visual thermal representation influenced the location where they grasped. Some participants reported *"I used the handle because I did not want to get burnt or my hand to be too cold."* [P34] or *"I used the handle for hot content to avoid being burnt"* [P28] while other participants reported *"Not necessarily as I wasn't too concerned about burning my hand (because it is robotic [sic]), therefore, didn't matter how I grasped the mug "* [P03].

Some additional comments included: *"With real hand, I almost expected the mugs to have different weights with respect to the amount of liquid in each mug. This was to a greater extent than the abstract hand."* [P11], *"I felt that the simulation with the human hand made me feel the need to be careful in case I burnt myself a lot more than when using the robot hand simulation."*[P24], *"I first thought* *that it was my hand and I realised that it was not only by looking at the short nails."*[P47]. Overall, participants reported that mug content and thermal cues have a stronger impact on their grasp choice and location than other environmental cues such as changes in the environment or in the hand representation which was also found in the results of this study. The responses can be found in Appendix B.

7.8 Discussion and Conclusions

Results in this study showed that even when haptic feedback is missing, subjects change their grasping interaction behaviour based on visual cues for thermal haptic feedback only. This is consistent with VR literature, where the use of pseudo-feedback has shown to influence user perception (Nunez, Zenner, Steinicke, Daiber, & Krüger, 2022), subjects perceiving different haptic sensations based on visual cues only (Kawabe, Ujitoko, & Yokosaka, 2022). This was the case in the results of this study, where participants predominantly grasped the handle of the mug for the hot and cold condition, while grasping the body of the mug for the empty condition, even though there was no real consequence at the end of the grasping task, regardless of the grasping location (e.g. the user cannot get burned if grasping the hot part of the mug). The reason behind this might be high levels of presence in VR which might have influenced subjects to naturally behave as they would in a real environment. This is consistent with VR literature showing that using visual cues to substitute haptic feedback is associated with increased feelings of presence, immersion and enjoyment (Rietzler et al., 2018).

The grasped location was essentially the metric that was directly influenced by thermal cues, however results showed there is a strong link between grasped location, GAp and grasp labels in VR. Grasping instances presenting a small GAp

were predominantly located around the handle of the virtual mug, showing the influence of the size of the grasped location on the grasp pattern when interacting with virtual objects. This is in alignment with findings in Chapter 6 and grasping real objects, where the aperture of the grasp is primarily influenced by the size of the grasped location (Feix et al., 2014b). Similarly, results highlighted that the shape and size of the grasped object (or part of the object) also influenced the grasp type chosen. Therefore, the hypothesis of H_1**: The thermal visual cues of the object have no effect on the grasp metrics** has is rejected as all grasp metrics showed differences between thermal visual cues.

However, thermal cues showed significant differences in terms of grasping metrics for all environmental conditions and hand representations except contextual objects condition for abstract hand. This is interesting since the basic and contextual objects conditions present the same mug (yellow mug with thermal conditions). This might have been due to the environment factors (additional objects and window view) influencing users to focus less on the mug with the thermal visual cues and therefore led to higher variability in grasp metrics. Another factor that might have influenced this effect is the hand representation, users feeling less connected with the abstract hand and therefore not being as careful with the thermal visual cues.

When comparing between hand representations, ES was generally smaller for the abstract hand than for the human hand, however, there were no significant statistical differences between hand representations for grasp data. Moreover, the IPQ scores did not present statistically significant differences in perceived presence between hand representations. This finding is contrary to popular literature (Schwind et al., 2017), and therefore the null hypothesis H_2**: The visual representation of the hand has no effect on the grasp metrics** has failed to be rejected. However, during the interview and post-questionnaire within the study, participants emphasised higher levels of attachment to the human hand.

7.8.1 Influences on the VR Taxonomy

This work aimed to explore whether grasp metrics are influenced by thermal visual cues in VR. Complementary to work in Chapter 6 which presented the first VR Taxonomy of Grasp Types based on object shape where users have shown to use familiar behaviours from reality when grasping virtual objects (e.g. wrapping their hand to match the size and shape of the object, with the accuracy limitations imposed by VR current hardware), this work also showed that users translate real behaviours in VR grasping by adapting their grasp location and inherently grasped aperture and grasp labels when thermal cues are shown on the virtual object. These findings therefore show that the original taxonomy could be further enhanced via thermal representation on virtual objects, which in this context showed to introduce changes in grasping approach.

For example, in Chapter 6, the Equant object category (which also contains the Mug object used in this experiment) has shown to be predominantly grasped using Large Diameter and Power Sphere grasp types. In this experiment, the mug was still grasped using Large Diameter and Power Sphere, however due to the effect of thermal visual cues, these grasping patterns were changed in hot and cold conditions, introducing a prevalence for Small Diameter and Precision grasps in general. This is due to users predominantly grasping the handle of the mug in the hot and cold conditions and predominantly grasping the body of the mug in the empty condition.

While a taxonomy of grasp types for thermal visual cues can not be fully adapted based on the findings in this experiment, due to the nature of the study only exploring one object and not representing objects from Equant, Prolate, Oblate and

Bladed, this work shows that the taxonomy might be changed when thermal visual cues are present, especially in the number and types of grasps associated with an object category, particularly when objects can be grasped on various locations which support the thermal context in which the objects are being used in (handles, protruding parts). However, future work needs to evaluate how other representations of thermal information on virtual objects change the current taxonomy of virtual grasps.

While this chapter provided an overview of the effect of pseudo-haptics on virtual grasping, it also emphasized the connection between grasp metrics in VR which is important for defining and improving grasp taxonomies for virtual objects. While GAp was under and overestimated for different grasped locations, subjects still adjusted their hand opening for every grasp location. With this adjustment they also changed the grasp type chosen suggesting that a more detailed correlation can be made between where the object is grasped and hand posture (grasp type), which will be explored in more detail in the next chapter, where a more complete iteration of the VR Taxonomy of Grasp Types is presented.

8 | Virtual Object Categorisation

This work was published in the proceedings of the 2021 IEEE Virtual Reality and 3D User Interfaces (VR) as "Freehand grasping: An analysis of grasping for docking tasks in virtual reality" (Blaga et al., 2021a) and in the proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology as "Virtual Object Categorisation Methods: Towards a Richer Understanding of Object Grasping for Virtual Reality" (Blaga et al., 2021c).

8.1 Introduction

Object categorisation is a milestone in the human evolutionary history, by which people can explore and understand the world faster and better (Guan & Zhang, 2016). It supports us in grouping elements together based on similar properties or attributes which can improve learning, prediction, decision making, language and interaction. Research into computer vision (Leibe & Schiele, 2003) has for many years developed on the ability to use categorisation to improve task goals, namely in artificial intelligence for behaviour characterisation (Prange, Barz, & Sonntag, 2018), text recognition (Kamal & Sultana, 2012), gesture analysis (Hummels & Stappers, 1998) and robot grasping (Kerzel et al., 2017), each leveraging the humans' ability to form and recognise categories of objects and define frameworks for computer-based categorisation. Commonly, these categorisation frameworks focused on specific attributes from human perception studies, one being Zingg's (Zingg, 1935) theory of using object dimension ratio to categorise objects based on shape, as described in Chapter 6. Other theories have focused on representing object components as arrangements of simple convex and concave parts (Biederman, Cooper, Hummel, & Fiser, 1993) or using wider object properties for categorisation such as auditory feedback (Guan & Zhang, 2016). Over recent years categorisation methods have supported researchers in uncovering a richer understanding of how humans interact with objects within the real world and how to develop improved HCI and computer vision systems. While Chapter 6 used Zingg's (Zingg, 1935) categorisation method to group objects based on their shape and develop the first VR Taxonomy of Grasp Types, other categorisation methods have not been explored for grasping in VR. To develop a more complete grasp taxonomy, this chapter presents categorisation methods that take into consideration object characteristics such as stability and protruding objects.

When analysing task influence in AR, researchers found that different tasks such as pointing, reaching, tilting (Mousavi Hondori, Khademi, Dodakian, Cramer, & Lopes, 2013), moving, rotating and scaling (Bai, Gao, El-Sana, & Billinghurst, 2013) require different interaction approaches when manipulating virtual objects (Piumsomboon et al., 2013). Task has shown to influence grasping patterns in real environments (Feix et al., 2014a), however, Chapter 6 showed that simple translate tasks did not influence grasping approach in VR. Yet, tasks that involve both translation and rotation have not been analysed for grasping patterns in VR. To address this, work in this chapter presents an elicitation study to define grasp patterns for freehand virtual object manipulation docking tasks in VR. Across 39 participants, 3 conditions and 16 virtual objects, this study reports on grasp metrics following the methodology described in Chapter 4.

The chapter is structured as follows: Section 8.2 presents a literature review of existing work on object categorisation methods and docking tasks in VR; Section 8.3 presents the virtual object categorisation with a detailed overview of methods proposed and the categorisation experiment; Section 8.4 presents the elicitation experiment design with a detailed methodology being presented for the task, hand representation and virtual environment; Section 8.5 presents the protocol for the user elicitation study as detailed in Chapter 4. Section 8.6 presents the metrics used for analysing grasping patterns; Section 8.7 presents the proposed hypotheses; Section 8.8 presents the methodology used for data analysis; Section 8.9 presents the results where grasp dimension and grasp labels are analysed and structured in a more complete VR Taxonomy of Grasp Types. Section 8.10 presents discussion and conclusions.

8.2 Background

8.2.1 Categorisation Methods

Categorisation methods have been previously used in research for grouping attributes and defining classes of syntactic and semantic features, and further used for automated characterisation (Prange et al., 2018). This then allowed the development of categorisation methods showing that organising and processing data in a simple, uniform and structural form hugely benefits development of interactive systems (Hummels & Stappers, 1998; Song, Qin, & Zhang, 2016).

Object classification based on shape has been investigated for a long time, with geon theory showing how objects can be represented as an arrangement of simple convex and concave parts (Biederman et al., 1993). However, since it has been shown that properties such as material, hardness, weight and roughness can only be detected or extracted by non-visual modalities, Guan and Zhang (Guan & Zhang, 2016) proposed a novel system of multimodal object recognition and categorisation by performing interactive behaviours, using both auditory and visual modality. Similar methods have been used for categorising real objects for understanding human grasping. Derbyshire et al. (Derbyshire, Ellis, & Tucker, 2006) performed an object categorisation study where they asked participants to categorise objects in two categories: naturally formed and manufactured and analysed reach to grasp actions for these object categories. Further, to group attributes of haptic materials, Kerzel et al. (Kerzel et al., 2017) presented a framework for haptic material classification based on an adaptation of human haptic explanatory procedures executed by a robot arm with an optical force sensor, showing that categorising haptic material is important for successful robot grasping. While categorisation methods have been highly explored for real grasping patterns, they have not been fully explored for virtual grasping yet. Chapter 6 used Zingg's (Zingg, 1935) categorisation method to analyse how grasping approach changes based on object shape. However, this methodology is limited, only covering shape and form, without considering other object characteristics such as roundness and complexity. The current VR Taxonomy of Grasp Types presents grasp patterns for object categories based on Zingg's (Zingg, 1935) method, however considering the above, grasping patterns for other categorisation methods need to be evaluated for a more complete taxonomy of grasp types in VR.

8.2.2 Docking Tasks in VR

Target acquisition is one of the most elementary interactions in 3D environments (El-Shimy, Marentakis, & Cooperstock, 2009), with researchers using variations of tasks such as placement, orientation and docking to assess aspects of virtual object manipulation (Martinet et al., 2010; Englmeier, Dörner, Butz, & Höllerer, 2020). Docking is the task that combines point or volume matching with orientation matching, also known as the simplest 6 Degrees of Freedom (DOF) task (Boritz & Booth, 1998). It is naturally performed by the human hand in everyday tasks for manipulating real objects, and therefore is a fundamental task for developing and assessing interaction metrics under different experimental conditions in areas such as robotics (Muse, Weber, & Wermter, 2006) and VR environments (Vuibert, Stuerzlinger, & Cooperstock, 2015). Vuibert et al. (Vuibert et al., 2015) used a docking task to assess gross motion and compare different interaction methods based on accuracy and completion time. Bai et al. (Bai, Nassani, Ens, & Billinghurst, 2017) implemented rotation and translation tasks for mid-air 3D object manipulation and proposed a comprehensive repertoire of 3D manipulation operations. Boritz et al. (Boritz & Booth, 1998) created placement and orientation tasks to assess different visual feedback modes and showed that target position and orientation on a docking task have a significant effect upon user performance in VR. Chapter 6 explored grasping patterns for simple translation tasks, and showed no statistical significant differences in grasping patterns between conditions. However, docking tasks and their influence on grasp metrics in VR has not been explored before. To address this, this chapter aims to explore docking tasks in VR and how they influence grasping metrics.

8.3 Virtual Object Categorisation

As presented in Chapter 6, previous work looking at grasping patterns in real environments (Feix et al., 2014b) used Zingg's methodology to categorise objects for assessing correlations between object manipulation metrics and grasping approach. While Chapter 6 showed that this categorisation method was useful in identifying grasping pattern changes between object categories, findings also showed higher variability in grasp choice for objects with more complex geometries. Chapter 7 showed that this might be connected to irregular shapes presenting more graspable locations, however, categorisation methods focused on object geometries and their influence on grasping patterns in VR have not been explored yet. Moreover, Chapter 6 also showed a change in grasping patterns (especially around grasp apertures) for virtual objects that present a level of roundness around certain axes (cylindrical, spherical shape), however, categorisation methods focused on object roundness/stability and their influence on grasping patterns in VR have not been explored yet. Therefore, to develop a more complete VR Taxonomy of Grasp Types that links object types to grasp types, this chapter explores grasping patterns for the following categorisation methods: Zingg's (Zingg, 1935), Virtual Object Equilibrium (VOE) and Parts.

8.3.1 Methods

Zingg's Method: As described in Chapter 6, Zingg's (Zingg, 1935) methodology categorises objects based on their shape and the three dimensions that indicate the volume of geometric bodies. Zingg (Zingg, 1935) defined A as the longest dimension of an object, C the shortest and B the remaining dimension. He defined a constant R to describe the relationship between dimensions and categorise the object; determining that the value at which one typically regards two axes to be different is about R = 3/2 (Zingg, 1935). Based on these parameters, four shape categories were defined as part of Zingg's categorisation framework: *Equant*, *Prolate*, *Oblate* and *Bladed* (see Figure 8.1)

VOEquilibrium Method: While Zingg's methodology has been previously used in grasping research, it does not take into account object roundness and stability, which might influence grasping approach (Howard & Kumar, 1996) (i. e. if the object is unstable and very likely to be rolled over if touched incorrectly, users might pick a different grasp type compared to a stable object). To address this, Szabo et al. (Szabo & Domokos, 2010) proposed a classification system which involves counting static equilibria, using a faster and easier framework than the classical Zingg method (Zingg, 1935). They define static equilibria as points of the surface where the object is at rest when placed on a horizontal, frictionless support surface. Moreover, grasping stability/equilibria has been investigated before (Howard & Kumar, 1996), showing that contact bodies (which are directly linked to the shape of the object) influence the stability of a grasp, showing the need for taking this into consideration when categorising virtual objects for understanding grasping patterns in VR. Therefore, this section introduces the VOEquilibrium (Virtual Object Equilibrium) method for object categorisation, which focuses on equilibrium points of virtual objects. As part of this method, this thesis presents two categories of virtual objects: Stable and Unstable. As shown in Figure 8.1, Stable category contains objects that are at rest when placed on a horizontal surface and would not be easily moved if incorrectly touched, while *Unstable* category contains objects that are easily moved if incorrectly touched or that do not automatically return to initial position of rest after a small perturbation.

Parts Method: Zingg's methodology focuses on object dimensions and overall shape, however, it does not take into account protruding objects such as handles which has shown to influence grasping patterns in Chapter 6 and 7. VOEquilibrium methodology focuses on object stability (equilibrium points), however, it does not take into account shape and object geometry. Geon theory (Biederman et al., 1993) addresses object geometry components and introduces the concept of breaking objects down in geons, to then categorise them based on the number of components (or geons). Inspired by geon theory, this thesis introduces Parts categorisation method for virtual objects, which includes two subcategories that focus on the number of graspable components, to allow a clear analysis of how object complexity influence grasping approach in VR (for example, if a mug has a handle, users could either grasp the body of the mug or the handle as shown in Chapter 7, which allows different grasping approaches for the same virtual object). Therefore, using this method, objects are categorised based on the number of parts in: One-Part and Multiple-Part. As described in Figure 8.1, One-part objects are composed of one geometric graspable component only, while Multiple-part objects are composed of more than one geometric shape.

8.3.2 Virtual Objects

The virtual objects used for this experiment were the 16 objects selected from the *Yale-Carnegie Mellon University-Berkeley Object and Model Set* (Calli et al., 2015) following the methodology detailed in Chapter 4 (Table 4.2). These objects were grouped in the four Zingg's (Zingg, 1935) categories in Chapter 6, however for grouping the virtual objects in VOE and Parts categories, a user categorisation study was conducted.

8.3.2.1 Categorisation Raters

A total of 15 raters (7 male and 8 female) from a population of university students and staff members volunteered to take part in this study. Participants ranged in age from 19 to 45 (M = 31.66, SD = 13.01). All participants completed the full categorisation for all objects.

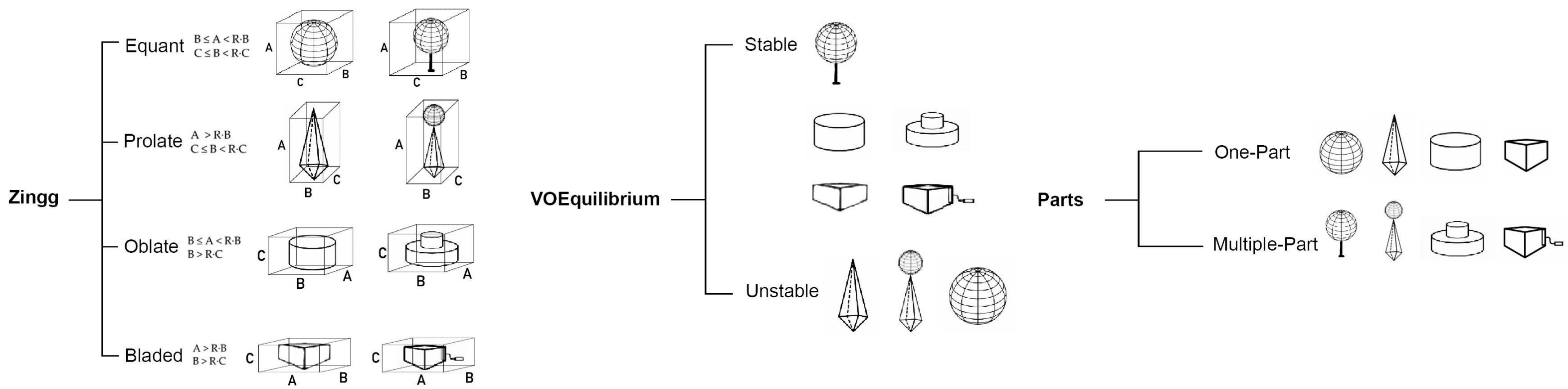


Figure 8.1: Categorisation methodologies discussed in this study along with 8 example abstract objects categorised for each methodology. Zingg's (Zingg, 1935) methodology shows the four categories: *Equant*, *Prolate*, *Oblate* and *Bladed* along with example abstract objects and their dimensions. VOEquilibrium methodology shows the two categories *Stable* and *Unstable* with example abstract objects. Parts shows the subcategories: *One-Part* and *Multiple-Part* along with example abstract objects.

8.3.3 Protocol

The categorisation process was conducted online, complying with COVID-19 regulations and guidelines and underwent a formal ethical review from the university ethical review committee. Raters were asked to first complete a training session, where each categorisation method (VOEquilibrium and Parts) was explained in terms of theoretical aspects with example images as shown in Figure 8.1. Further, raters completed a questionnaire where they were asked to categorise the 16 virtual objects used in this experiment. Results for Zingg's (Zingg, 1935) categorisation method are extracted from work presented in Chapter 6. For VOEquilibrium and Parts categorisation methods, raters divided the 16 objects in the categories presented; For VOEquilibrium in Stable and Unstable and for Parts in One-Part and Multiple-Part.

8.3.3.1 Categorisation Agreement

Following Wobbrock et al. (Wobbrock et al., 2005) definition, the categorisation agreement score was defined as the agreement among the categories proposed by raters per object as shown in Table 8.1. The level of agreement per object was computed following the equation:

$$\frac{\sum_{r \epsilon R} \sum_{P_i \subseteq P_r} \left(\frac{P_i}{P_r}\right)^2}{|R|} \tag{8.1}$$

Where *r* is an object in the set of all available objects *R*; P_r is the set of category proposals for object *r* and P_i is a subset of identical categories for P_r as in (Wobbrock et al., 2005, 2009). For VOE categorisation, 12 objects showed an agreement equal or higher than 73% and 4 objects showed an agreement between 57%-73%. For Parts categorisation, 6 objects showed an agreement equal or higher than 73% and 10 objects showed an agreement between 52%-73%.

8.3.4 Categorisation Results

The virtual objects were categorised following Zingg's methodology in Chapter 6 (four objects in Equant, four objects in Prolate, four objects in Oblate and four objects in Bladed). This is shown in Table 8.1 together with the results from the categorisation study presented in this work, with the objects being again categorised following the two novel methods proposed: VOE and Parts. Following the VOE methodology, virtual objects were categorised as follows: Brick, Mug, Cracker Box, Clamp, Cleanser Bottle, Mustard, Lego, Hammer, Gelatine Box, Meat Can, Sponge and Scissors were categorised as Stable, while Banana, Marker, Orange and Spoon were categorised as Unstable. Following the Parts methodology, virtual objects were categorised as follows: Brick, Banana, Cracker Box, Lego, Orange, Gelatine Box, Meat Can and Sponge were categorised as One-Part while Mug, Marker, Clamp, Cleanser Bottle, Mustard, Hammer, Spoon and Scissors were categorised as Multiple-Part.

Object	Zingg		VOE		Parts		Object	Zingg		VOE		Parts	
	C	A	C	A	C	A		C	A	C	A	C	A
	Equant	N/A	Stable	1	One-Part	0.73		Equant	N/A	Stable	0.85	One-Part	0.57
	Equant	N/A	Stable	0.64	Multiple-Part	0.52		Equant	N/A	Unstable	1	One-Part	0.85
	Prolate	N/A	Unstable	0.73	One-Part	0.85		Prolate	N/A	Stable	0.57	Multiple-Part	0.73
	Prolate	N/A	Unstable	0.64	Multiple-Part	0.52		Prolate	N/A	Unstable	0.64	Multiple-Part	0.57
	Oblate	N/A	Stable	0.85	One-Part	0.57		Oblate	N/A	Stable	0.73	One-Part	0.57
	Oblate	N/A	Stable	0.73	Multiple-Part	0.64		Oblate	N/A	Stable	1	One-Part	0.52
	Bladed	N/A	Stable	0.73	Multiple-Part	0.57		Bladed	N/A	Stable	0.85	One-Part	0.73
	Bladed	N/A	Stable	0.73	Multiple-Part	0.57		Bladed	N/A	Stable	1	Multiple-Part	0.85

Table 8.1: Results showing the categorisation of virtual objects. Column Object shows the visual representation of the virtual objects used in this study. For each categorisation method explored, there are two subcolumns: C, showing the category and A showing the agreement score. Column Zingg shows the categories from Zingg's methodology (Zingg, 1935) with agreement score N/A since the results are taken from calculating the equations and not during the user experiment. Column VOE shows the categorisation results from VOEquilibrium methodology. Column Parts shows the assigned category based on Parts methodology.

8.4 Docking Experiment Design

To analyse how object categorisation influences grasping interaction patterns, a user experiment was designed following the methodology detailed in Chapter 4. This section provides a detailed overview of the apparatus, docking task, environment and 3D hand model used for interaction as detailed in Table 4.3.

8.4.1 Apparatus

As in Chapters 5-7, a custom experimental framework was built using the Oculus Rift DK2 VR Head Mounted Display (HMD) and the Leap Motion tracking device as detailed in Chapter 4. The Leap Motion was mounted on the front of the HMD facing the participants' hands, to facilitate hand interaction. A web camera was also used for recording participants' hand during interaction. The system was developed using C#, Unity 2018.2 and Leap Motion 4.0 SDK.

8.4.2 Task

Task intention has shown to influence grasp patterns when interacting with real objects, presenting different grasp choices for the same object based on the task to be performed (Feix et al., 2014b). Moreover, VR systems where natural interactions have an impact on overall performance are task-oriented (Ma et al., 2010). When manipulating virtual 3D content in VR, placement, orientation and docking (6 DOF) are fundamental tasks for which the choice of interaction gestures is critical for usability and performance (Martinet et al., 2010). Therefore, to understand the grasp choice in VR and how this changes for various tasks, a docking task was designed for 3D object manipulation, namely placement and orientation of a virtual object to a target position and rotation as shown in Figure 8.4.

Virtual Object Rotation: Research shows that virtual rotation tasks are more challenging than real rotation tasks (Ware & Arsenault, 2004). To assess if this has an influence on grasping patterns in VR, the virtual objects were rotated on the horizontal plane (azimuth) by a value of 30°, 60°and 90 °, with objects being ergonomically reachable by the hand. To avoid confusion regarding the direction of rotation, the target for the docking task was always facing frontwards.

Target Categories: Aiming to analyse how rotation and translation influence grasping patterns in a realistic task, the task was designed to show target categories where users had to place the virtual objects, similarly to cleaning up a table in daily life. The target categories used were: Tools (Figure 8.2 a), Groceries (Figure 8.2 b) and Fruits (Figure 8.2 c).

Figure 8.2: Target categories *Tools, Groceries, Fruits* showing the object targets categorised by their daily usage.

8.4.3 Conditions

The study consisted of three conditions where the position of the target objects and the rotation of the interaction objects change as follows: **[DC1]**: Tools on the right, Groceries in the centre and Fruits on the left, with 30°rotation (Figure 8.3(a)); **[DC2]**: Fruits on the left, Tools in the centre and Groceries on the right, with 60°rotation (Figure 8.3(b)) and **[DC3]**: Groceries on the left, Fruits in the centre and Tools on the right, with 90°rotation (Fig 8.3(c)). Virtual objects were randomly positioned on the virtual table, each of the objects having a corresponding target position and rotation (highlighted in green in Figure 8.3). At the end of the docking task all objects had to be in the target position and rotation.

(a) (b) (c)

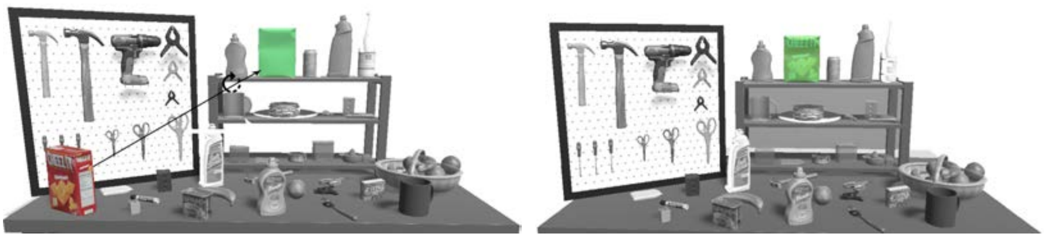
Figure 8.3: Virtual environment conditions. (a) Docking Condition 1 [DC1] showing *Tools* at the left, *Groceries* in the centre and *Fruits* on the right, with a 30°rotation; (b) Docking Condition 2 [DC2] showing *Fruits* on the left, *Tools* in the centre and *Groceries* on the right, with a 60°rotation; (c) Docking Condition 3 [DC3] showing *Groceries* on the left, *Fruits* in the centre and *Tools* on the right, with a 90°rotation.

8.4.4 Hand Representation

Chapter 7 explored the effect of hand representation on grasping patterns in VR. While no significant differences were found between abstract and human representations, participants showed a preference for the human hand representation. Therefore, this experiment used the human hand 3D model to represent participants' hand during grasping interaction.

8.4.5 Environment

In connection to the methodology in Chapters 5-7, the experiment was conducted in a controlled environment under laboratory conditions. The test room was lit by a 2700k (*warm white*) fluorescent light with no external light source. The virtual environment consisted of a virtual table and the target categories (tools, groceries and fruits) presenting the target objects. For each task, the virtual objects appeared on the virtual table, randomly arranged at 10 cm distances from each other.



(a) Docking task before completion (b) Docking task after completion

Figure 8.4: Example of one of docking task for the *Cracker Box* virtual object. (a) Docking before completion shows the task before the user grasps it and translate + rotate it to the target position (highlighted in green); (b) Docking task after completion shows the task after the target was translated and rotation to the target position (overlaying the green area).

8.4.6 Participants

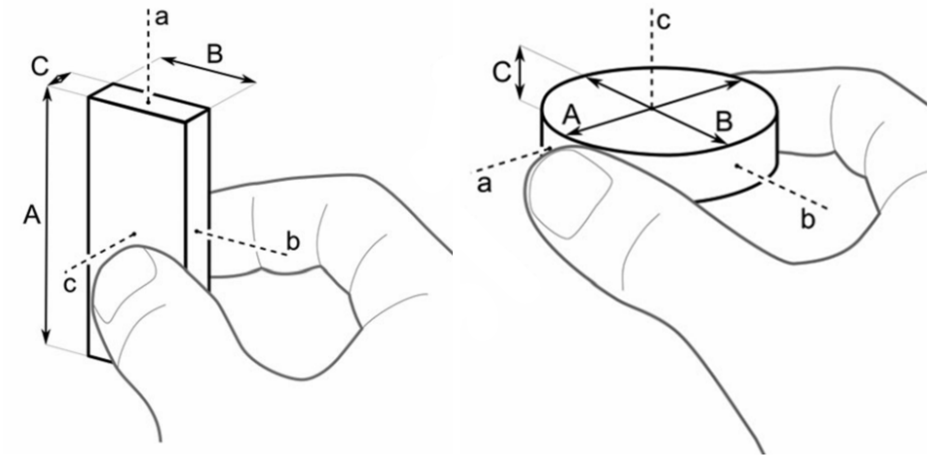
A total of 39 participants (25 male and 14 female) from a population of university students and staff members volunteered to take part in this study. Participants' age ranged from 19 to 47 (M = 27.69, SD = 6.66). All participants were right-handed, to ensure they interacted with the virtual objects under the same conditions. All participants performed the full study. As in Chapters 5-7, visual acuity of participants was measured using a Snellen chart. Each participant was also required to pass an Ishihara test to check for colour blindness. Participants with colour blindness and/or non corrected visual acuity of < 0.80 (where 20/20 is 1.0) were not included in this study. Participants were asked to self-assess their level of experience with VR systems and gesture recognition systems. Regarding VR systems, 21 reported being novice to the technology, 16 reported having an average level of experience and 2 self-labelled themselves as experts. Regarding gesture-recognition systems, 27 reported being novice to the technology, 10 reported an average level of experience and 2 self-labelled themselves as experts. Participants were not compensated.

8.5 Protocol

Pre-test: Prior to the study, participants were given a consent form where the test protocol and main aim of the study was described. Additionally, participants completed a pre-test questionnaire enquiring about their background level of experience with VR systems and hand recognition sensors.

Training: Participants underwent initial hand interaction and task training to familiarise themselves with the VR environment. The training task was a representative version of the tasks in the study (grasp, translate and rotate a cube in the 3D space).

Test: Each participant completed 48 grasps (3 conditions $\times$ 16 objects), with a total of 1872 grasps recorded during the study (48 grasps $\times$ 39 participants). Participants were seated. The virtual objects were randomly positioned on the virtual table for each condition and were consistent across participants. A Wizard of Oz approach was followed and participants were instructed to grasp the virtual objects the way they felt most intuitive, notifying the test instructor when they were happy with their grasp, as in Chapters 5-7. When they were happy with the position and rotation, they released the object and moved to grasping another object from the virtual table. Users were allowed to grasp the objects in any order. One condition was complete when all objects from the virtual table were arranged at their target location and rotation.



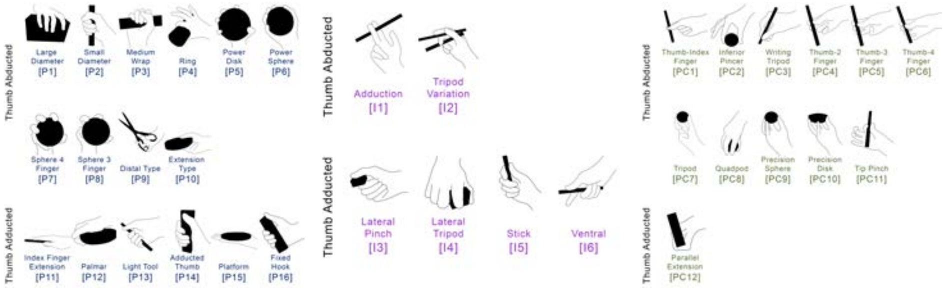
(a) Grasp Dimension C (b) Grasp Dimension A/B

Figure 8.5: Grasped dimension examples as defined by Feix et al. (Feix et al., 2014b).

Post-test: After all tasks were completed, participants were asked to complete a post-test questionnaire to describe their perspective on grasping the virtual objects.

8.6 Metrics

Grasp Dimension: As detailed in Chapter 4, Feix et al. (Feix et al., 2014b) defined grasped location as the part of the object that lies between the fingers when grasped. By using the object axes (A, B and C) as defined during Zingg's categorisation method (Zingg, 1935), grasped dimension is analysed to indicate which axes best determine the hand opening. Figure 8.5 shows examples of how object dimensions determine the grasped dimension. In Figure 8.5 a) the object is grasped along the shortest dimension (dimension C) while in Figure 8.5 b) the object is grasped along dimension A/B, meaning that both A and B dimensions determine the hand opening. Chapter 7 explored grasped location, to analyse where subjects grasp the virtual mug, however grasped location is unique for each indi-



(a) Power Grasps (b) Intermediate Grasps (c) Precision Grasps

Figure 8.6: Grasp types from the real grasp taxonomy (Feix et al., 2009)

vidual object (e.g. mug would have handle, body and top while bottle would have body and lid). To allow a more generalisable analysis on where virtual objects are grasped based on their characteristics, grasp dimension is analysed for every grasp instance during labelling as presented in Chapter 4.

Grasp Labels: As in Chapters 5-7, the grasp categories are Power, Intermediate and Precision. Power grasps are linked to stability and security (Figure 8.6 a). Intermediate grasps present elements of Power and Precision roughly in the same proportion, enabling a finer representation of grasp types (Bullock et al., 2013) (Figure 8.6 b). Precision grasps present grasps where the object is commonly held between the fingertips (Figure 8.6 c).

8.7 Hypotheses

Chapter 6 showed that object shape influences grasping patterns in VR, which has been also shown in immersive virtual object grasping (Al-Kalbani et al., 2016a). Therefore, the following hypothesis is proposed:

H_1**:** Categorisation based on object characteristics influences grasping patterns in VR.

Chapter 6 showed that simple translate tasks did not influence grasping patterns in VR, however the influence of docking tasks on grasp metrics in VR is unknown, therefore the following null hypothesis is proposed:

H_2**:** Grasping patterns in VR are not influenced by rotation and translation docking tasks.

8.8 Data Analysis

The Shapiro-Wilk (Shapiro & Wilk, 1965) normality test found the data to be not normally distributed. Statistical significance between conditions and object categories in terms of grasping patterns, where the dependent variable (Grasp Dimension and Grasp Category) is nominal categorical (A, B, C and Power, Intermediate and Precision), contingency tables were created and analysed for significance using a Chi-Squared Test of Independence with 95% Confidence Intervals, therefore, a p-value of less than .05 will indicate statistical significance. Cramer's V calculation for effect sizes was applied after verifying its assumptions that the variables under analysis are categorical. Results were interpreted following existing guidelines based on degrees of freedom.

8.9 Results

8.9.1 Grasp Dimension ($GDim$)

Users predominantly grasped virtual objects presented in this study along the C dimension (66.93%, N = 1253), followed by dimension B (32.47%, N = 608) and dimension A (0.58%, N = 11). To understand the link between docking tasks, object categorisation and grasp dimension, statistical significance was tested for grasp dimension between docking conditions (DC1, DC2 and DC3) and object categories of the three methods explored (Zingg, VOEquilibrium and Parts) with 95% Confidence Intervals.

Docking Task: Statistical significance was found when comparing docking tasks in terms of grasp dimension (**χ^2 (4, N = 1872) = 13.79, p = .007*)** with small ES (Cramer's V = 0.06): [DC1] was generally grasped along dimension C (68.42%, N = 427) followed by dimension B (31.25%, N = 195) and dimension A (0.32%, N = 2). [DC2] was generally grasped along dimension C (70.67%, N = 441) followed by dimension B (28.52%, N = 178) and dimension A (0.81%, N = 5). [DC3] was generally grasped along dimension C (61.69%, N = 385) followed by dimension B (37.66%, N = 235) and dimension A (0.64%, N = 4). To understand how these patterns change based on object characteristics, statistical significance was tested between docking conditions for object categories of the three methods explored:

- Zingg's method: Statistical significance was found between docking tasks for Bladed category which for [DC1] was predominantly grasped around dimension B (51.92%, N = 81) followed by C (48.07%, N = 75) and A (0%); for [DC2] was predominantly grasped around dimension C (53.20%, N = 83), followed by B (45.51%, N = 71) and A (1.28%, N = 2); for [DC3] was predominantly grasped around dimension B (62.17%, N = 97), followed by C (36.53%, N = 57) and A (1.28%, N = 2): **χ^2 (6, N = 468) = 9.56, p = .041*** with medium ES (Cramer's V = 0.11).

  No significance was found for Equant category: [DC1] was grasped around dimension C (86.53%, N = 135), followed by B (13.46%, N = 21) and A (0%); [DC2] was grasped around dimension C (83.97%, N = 131), followed by B (16.02%, N = 25) and A (0%); [DC3] was grasped around dimension C (79.48%, N = 124), followed by B (20.51%, N = 32) and A (0%): χ^2 *(6,* *N = 468) = 3.30, p = .507*, Prolate category: [DC1] was grasped around dimension C (64.10%, N = 100), followed by B (35.89%, N = 56) and A (0%); [DC2] was grasped around dimension C (67.30%, N = 105), followed by B (32.69%, N = 51) and A (0%); [DC3] was grasped around dimension C (62.82%, N = 98), followed by B (36.53%, N = 57) and A (0.64%, N =1): χ^2 *(6, N = 468) = 9.31, p = .053* and Oblate category: [DC1] was grasped around dimension C (75%, N = 117), followed by B (23.71%, N = 37) and A (1.28%, N = 2); [DC2] was grasped around dimension C (78.20%, N = 122), followed by B (19.87%, N = 31) and A (1.28%, N = 2); [DC3] was grasped around dimension C (78.84%, N = 123), followed by B (20.51%, N = 32) and A (0.64%, N = 1): χ^2 *(6, N = 468) = 2.38, p = .665.*

- VOEquilibrium method: No statistical significance was found between docking tasks for Stable category: [DC1] was grasped around dimension C (63.03%, N = 295), followed by B (36.53%, N = 171) and A (0.42%, N = 2); [DC2] was grasped around dimension C (66.45%, N = 311), followed by B (32.47%, N = 152) and A (1.06%, N = 5); [DC3] was grasped around dimension C (59.61%, N = 279), followed by B (39.52%, N = 185) and A (0.85%, N = 4): χ^2 *(2, N = 1404) = 6.24, p = .181* and Unstable category: [DC1] was grasped around dimension C (84.61%, N = 132), followed by B (15.38%, N = 24) and A (0%); [DC2] was grasped around dimension C (83.33%, N = 130), followed by B (16.66%, N = 26) and A (0%); [DC3] was grasped around dimension C (78.84%, N = 123), followed by B (21.15%, N = 33) and A (0%): χ^2 *(2, N = 468) = 2.02, p = .731.*

- Parts method: No significance was found between docking tasks for Parts categories, One-Part: [DC1] was grasped around dimension C (83.33%, N = 260), followed by B (15.38%, N = 24) and A (0.64%, N = 2); [DC2] was grasped around dimension C (81.41%, N = 254), followed by B (17.62%,

Method	Category	DC1			DC2			DC3			Statistic
		A	B	C	A	B	C	A	B	C	
Zingg's	Equant	0	13%	86%	0	16%	83%	0	20%	79%	Stat = 3.30, p = .507
	Prolate	0	35%	64%	0	32%	67%	0	36%	62%	Stat = 9.31, p = .053
	Oblate	1%	23%	75%	1%	19%	78%	0	20%	78%	Stat = 2.38, p = .665
	Bladed	0	51%	48%	1%	45%	53%	1%	62%	36%	**Stat = 9.56, p = .041***
VOE	Stable	0	36%	63%	1%	32%	66%	0	39%	59%	Stat = 6.24, p = .181
	Unstable	0	15%	84%	0	16%	83%	0	21%	78%	Stat = 2.02, p = .731
Parts	One-Part	0	15%	83%	0	17%	81%	0	23%	76%	Stat = 7.52, p= .111
	Multiple-Part	0	46%	53%	0	39%	59%	0	47%	51%	Stat = 5.95, p = .202

Table 8.2: Grasp dimension results for each object categorisation method (with the corresponding categories) and each docking task explored in this study (DC1, DC2, DC3). Results are shown in percentages with statistical results reported for each object category.

N = 55) and A (0.96%, N = 3); [DC3] was grasped around dimension C (76.28%, N = 238), followed by B (23.39%, N = 73) and A (0.32%, N = 1): χ^2 *(2, N = 936) = 7.52, p = .111* and Multiple-Part: [DC1] was grasped around dimension C (53.52%, N = 167), followed by B (46.47%, N = 145) and A (0%); [DC2] was grasped around dimension C (59.93%, N = 187), followed by B (39.42%, N = 123) and A (0.64%, N = 2); [DC3] was grasped around dimension C (51.36%, N = 169), followed by B (47.72%, N = 157) and A (0.91%, N = 3): χ^2 *(2, N = 936) = 5.95, p = .202.*

Object Categorisation: Statistical significance was tested for overall object categories to understand the link between grasp dimension and object characteristics.

- Zingg's method: Statistical significance was found between object categories: Equant was predominantly grasped around dimension C (83.33%, N = 390), followed by B (16.66%, N = 78) and A (0%). Prolate was predominantly grasped around dimension C (64.74%, N = 303), followed by B (35.04%, N = 164) and A (0.21%, N = 1). Oblate was predominantly grasped around dimension C (77.35%, N = 362), followed by B (21.36%, N = 100) and A (1.28%, N = 6). Bladed was predominantly grasped around dimension B (53.20%, N = 249), followed by C (45.94%, N = 215) and A

Method	Category	A	B	C	Statistic
Zingg	Equant	0	16%	83%	**Stat = 171.45, p $\leq$ 0.001***
	Prolate	0	35%	64%	
	Oblate	1%	21%	77%	
	Bladed	0	53%	45%	
VOE	Stable	0	36%	63%	**Stat = 62.73, p $\leq$ 0.001***
	Unstable	0	17%	82%	
Parts	One-Part	0	19%	80%	**Stat = 140.53, p $\leq$ 0.001***
	Multi-Part	0	44%	54%	

Table 8.3: Grasp location (A, B, C) results for each object categorisation method and each category in percentages. Statistical results are shown for comparing against categories for each object categorisation method.

(0.85%, N = 157): χ^2 **(6, N = 1872) = 171.45, p < 0.001*** with large ES (Cramer's V = 0.42).

- VOEquilibrium method: Statistical significance was found between object categories: Stable was predominantly grasped around dimension C (63.03%, N = 885), followed by B (36.18%, N = 508) and A (0.78%, N = 11). Unstable was predominantly grasped around dimension C (82.26%, N = 385), followed by B (17.73%, N = 83) and A (0%):χ^2 **(2, N = 1872) = 62.73, p < 0.001*** with small ES (Cramer's V = 0.12).
- Parts method: Statistical significance was found between object categories: One-Part was predominantly grasped around dimension C (80.34%, N = 752), followed by B (19.01%, N = 178) and A (0.64%, N = 6). Multi-Part was predominantly grasped around dimension C (54.87%, N = 523), followed by B (44.59%, N = 425) and A (0.52%, N = 5):χ^2 **(2, N = 1872) = 140.53, p < 0.001*** with small ES (Cramer's V = 0.19).

Tables 8.4 and 8.5 show the split between dimension A, B and C used for grasping for every object category and condition explored in this experiment.

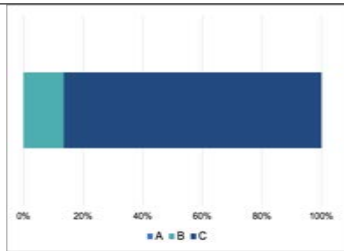
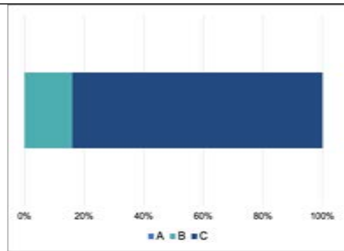
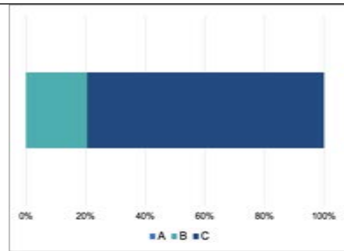
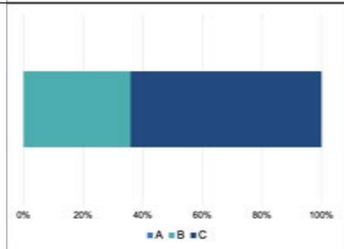
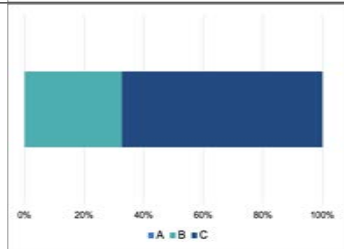
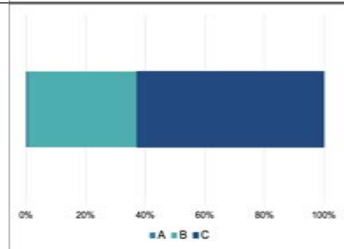
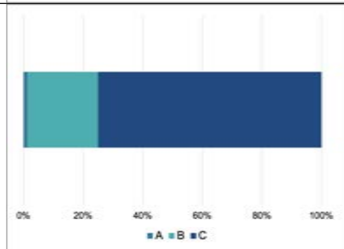
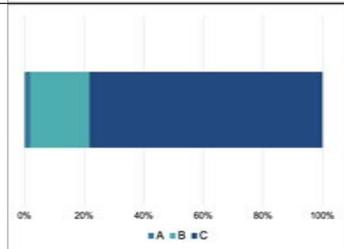
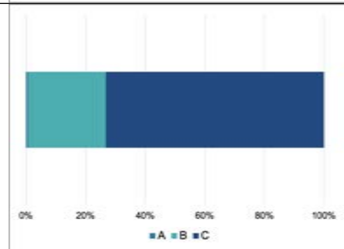
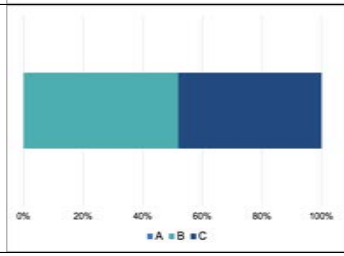
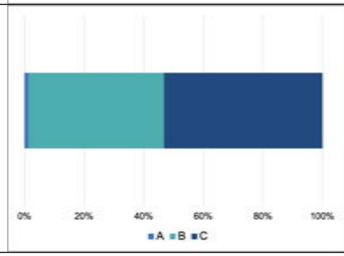
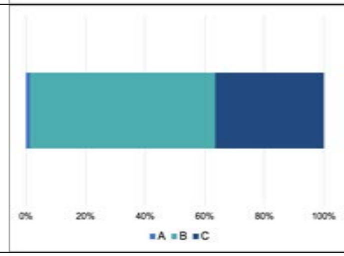
Category	$GDim$ - [DC1]	$GDim$ - [DC2]	$GDim$ - [DC3]
Equant			
Prolate			
Oblate			
Bladed			

Table 8.4: Grasp Dimension (A,B,C) for Zingg's object categories for every docking condition used in this experiment: DC1, DC2 and DC3.

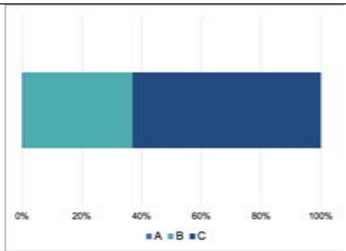
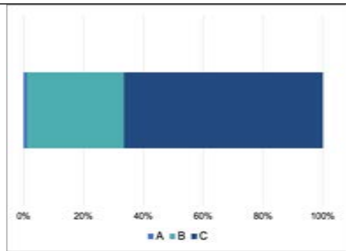
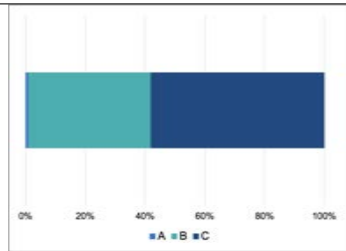
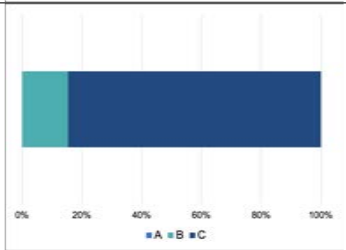
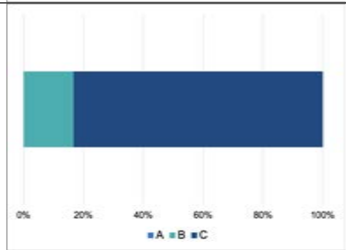
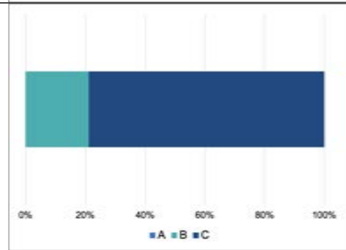
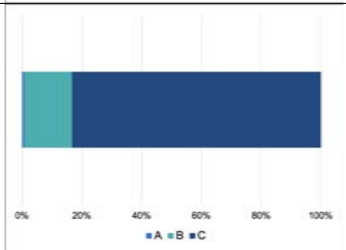
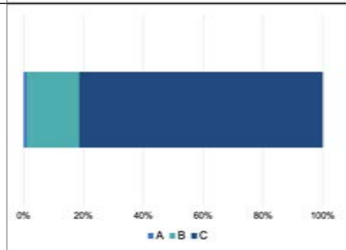
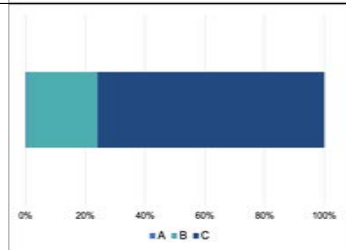
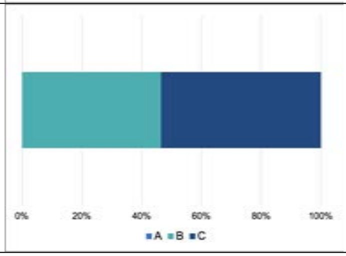
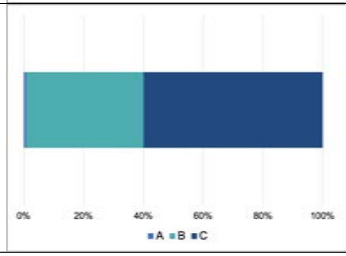
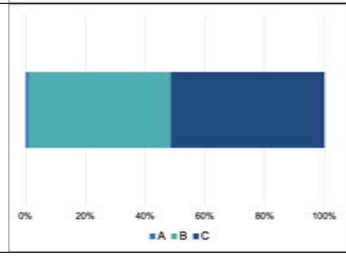
Category	$GDim$ - [DC1]	$GDim$ - [DC2]	$GDim$ - [DC3]
Stable			
Unstable			
One-Part			
Multi-Part			

Table 8.5: Grasp Dimension (A,B,C) for Stable, Unstable, One-Part and Multiple-Part object categories for every docking condition used in this experiment: DC1, DC2 and DC3.

8.9.1.1 Analysis - Grasp Dimension

Results showed similar patterns in terms of grasp dimension (Figure 8.5) between conditions under experiment for each object category, with objects being predominantly grasped along dimension C, the only exception being the Bladed category, which was predominantly grasped along dimension B. Real grasping literature analysing grasps of two machinists and two housekeepers showed that 94% of grasp instances collected were around the smallest dimension (dimension C), 4% around dimension B and 1% around dimension A (Feix et al., 2014b). While the prevalence for dimension C identified in this study is consistent with real grasping literature, in VR subjects grasped along dimension B more often, with $\approx$ 68% of total grasp instances being around C, 32% of total grasp instances being around B and 0.5% being around A.

When comparing between object categories, significant differences in terms of grasped dimension were found for all categorisation methods explored in this study. For Zingg's (Zingg, 1935) categorisation method, Equant and Oblate showed similar patterns in the split between B and C dimensions, while Prolate and Bladed showed a more balanced split between C and B dimensions. Prolate objects present long and narrow shapes, which in real grasping are generally grasped from the side (Feix et al., 2014b), which is consistent with findings in this experiment where the grasped dimensions, B and C represent the width and depth dimensions of the virtual objects. A similar pattern was found for Bladed objects, where the split between B and C was $\approx$ 50%. Bladed objects present a very small C dimension, compared to the other object dimensions, similar to a disk shape, and have shown to be predominantly grasped on dimension B in real environments (Feix et al., 2014b). This was evident in the results of this experiment where B dimension was chosen more frequently compared to the other object categories, however subjects still used dimension C for a significant number of grasp instances. While grasping really small objects might require additional precision and improved sensory acuity during a grasp (Napier, 1956) in real environments, as shown in Chapter 6 and 7, in VR users grasp objects larger and smaller than object size due to errors in object size estimations caused by the lack of haptic feedback. Therefore, this lack of haptic feedback allows users to grasp really small objects larger, without requiring more sensory acuity or mechanical support during their grasp, which might be why subjects also grasped these objects around dimension C.

For VOE categorisation method, Stable category showed an increased number of instances where the object was grasped around dimension B compared to Unstable category. Unstable objects are objects that generally require more precision when grasping due to their roundness and therefore proposing the risk of them rolling down the table if touched incorrectly. This prevalence for dimension C in Unstable objects might be due to users applying more precision when grasping these objects, due to learned actions from real grasping, however, this assumption is discussed in more detail in the following sections that focus on the type of grasps used for each object category. For Parts categorisation method, Multiple-Part showed a higher variability in grasped dimension than One-Part, which has been shown in real object grasping where irregular objects showed the largest variation in grasped dimension (Feix et al., 2014b).

Dimension A was very rarely grasped in both real grasping studies and results found in this experiment. Feix et al. (Feix et al., 2014b) motivated that dimension A is not generally grasped due to usually being larger than the human hand opening, therefore making grasping impossible along this dimension in real environments. In VR, grasping along this dimension is possible, considering that the hand can sink inside the object when the object dimension is larger than the hand

(a) Grasping on small dimension (b) Grasping on large dimension

Figure 8.7: Grasp examples showing a) grasp instance where the user mimics real grasping by wrapping the fingers around the object on dimension C and b) grasp instance where the user grasps along dimension B by sinking their hand inside the virtual object.

opening as shown in Figure 8.7. While there were some instances where subjects used this dimension for grasping, subjects generally chose smaller dimensions (B/C) when grasping, therefore mimicking real grasping behaviour in VR.

While real object grasping showed differences in grasping patterns based on task (Feix et al., 2014a), results in this study only showed significant differences between docking condition for Bladed objects, where the observed pattern is the prevalence of dimension B being chosen for grasping in [DC3] (90°rotation). While significant differences were not found for other object categories, this pattern for increased number of instances for dimension B in [DC3] can be observed for all categories under analysis. This might suggest that subjects changed the grasp dimension based on the target rotation and grasped in a way that supports a comfortable rotation of the object, however there is not enough evidence to support this claim in this experiment. The next section explores how docking tasks and object characteristics influenced the grasp types used for interaction.

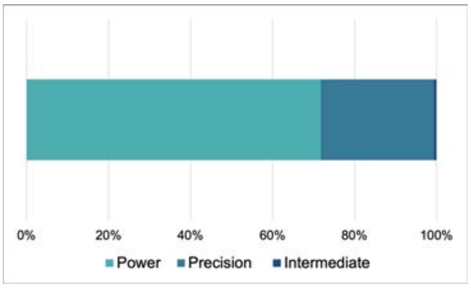
8.9.2 Grasp Labels

A total of 1872 grasps (39 participants x 16 objects x 3 conditions) were labelled during this user experiment, following the methodology presented in Chapter 4. No grasps were removed during the labelling process and all 1872 grasps were used in analysis. Cohen's Kappa was used to measure inter-rater reliability for labelling the collected grasps. Raters agreed in 87% of instances (Cohen's Kappa = 0.45) which based on existing guidelines is a moderate agreement, often achieved when subjectivity is involved in the process (S. Sun, 2011) and is a common agreement score for classification tasks (Feix et al., 2014c).

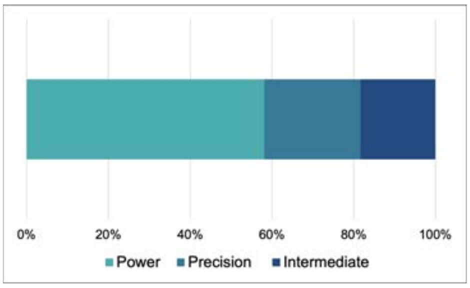
8.9.2.1 Grasp Category

Users predominantly grasped virtual objects presented in this study using a Power grasp (62.39%, N = 1168) followed by Precision (32.79%, N = 614) and Intermediate (4.80%, N = 90). To understand the link between docking tasks, object categorisation and grasp dimension, statistical significance was tested for grasp category between docking conditions (DC1, DC2 and DC3) and object categories of the three methods explored (Zingg, VOE and Parts) with 95% Confidence Intervals.

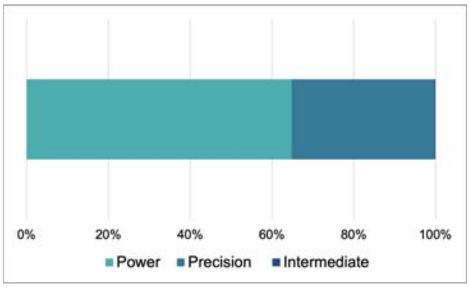
Docking Task: No significant statistical significance was found when comparing docking tasks in terms of grasp category (χ^2 (4, N = 1872) = 8.33, p = .081) with all conditions showing similar patterns in the use of grasp category. [DC1] patterns showed Power (59.93%, N = 374); Precision (34.93%, N = 218); Intermediate (5.12%, N = 32), [DC2] patterns showed Power (62.01%, N = 387); Precision (31.89%, N = 199); Intermediate (6.08%, N = 38) and [DC3] patterns showed Power (65.22%, N = 407); Precision (31.57%, N = 197); Intermediate (3.20%, N = 20).



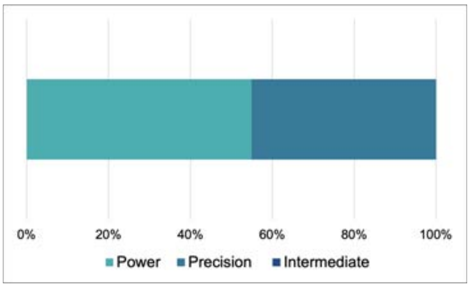
(a) Equant



(b) Prolate



(c) Oblate



(d) Bladed

Figure 8.8: The use of Power, Precision and Intermediate grasps for Zingg's (Zingg, 1935) object categories: Equant, Prolate, Oblate and Bladed

Object Categorisation: Statistical significance was tested for overall object categories to understand the link between grasp category and object characteristics.

- Zingg: Statistical significance was found between object categories: χ^2 **(6, N = 1872) = 285.01, p < 0.001*** with Equant being predominantly grasped using Power grasps (71.79%, N = 336), followed by Precision grasps (27.56%, N = 129) and Intermediate grasps (0.64%, N = 3); Prolate being predominantly grasped using Power grasps (58.11%, N = 272), followed by Precision grasps (23.50%, N = 110) and Intermediate grasps (18.37%, N = 86); Oblate being predominantly grasped using Power grasps (64.74%, N = 303), followed by Precision grasps (35.04%, N = 164) and Intermediate grasps (0.21%, N = 1); Bladed being predominantly grasped using Power grasps (54.91%, N = 257), followed by Precision grasps (45.08%,

Method	Category	Power	Precision	Intermediate	Statistic
Zingg	Equant	71	27	0	**Stat = 285.01, p < 0.001***
	Prolate	58%	23%	18%	
	Oblate	64%	35%	0	
	Bladed	54%	45%	0	
VOE	Stable	63%	36%	0	**Stat = 255.07, p < 0.001***
	Unstable	60%	21%	18%	
Parts	One-Part	66%	32%	0	**Stat = 77.17, p < 0.001***
	Multi-Part	57%	33%	9%	

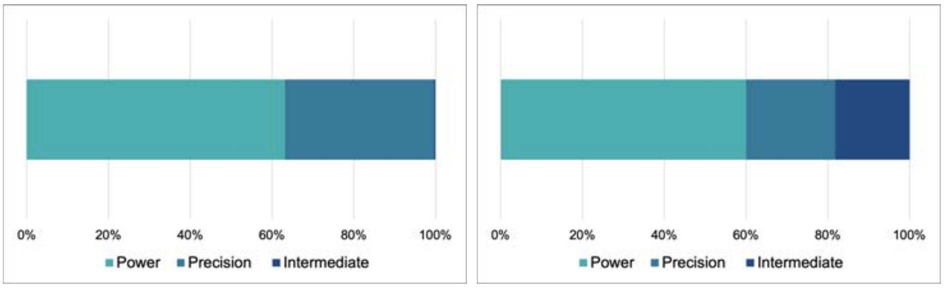
Table 8.6: Grasp category (Power, Precision and Intermediate) results for each object categorisation method and each category in percentages. Statistical results are shown for comparing against categories for each object categorisation method.

N = 211) and Intermediate grasps (0%).

- VOEquilibrium: Statistical significance was found between object categories: χ^2 **(2, N = 1872) = 255.07, p < 0.001*** with Stable being predominantly grasped using Power grasps (63.17%, N = 887), followed by Precision grasps (36.46%, N = 512) and Intermediate grasps (0.35%, N = 5) and Unstable being predominantly grasped using Power grasps (60.04%, N = 281), followed by Precision grasps (21.79%, N = 102) and Intermediate grasps (18.16%, N = 85).

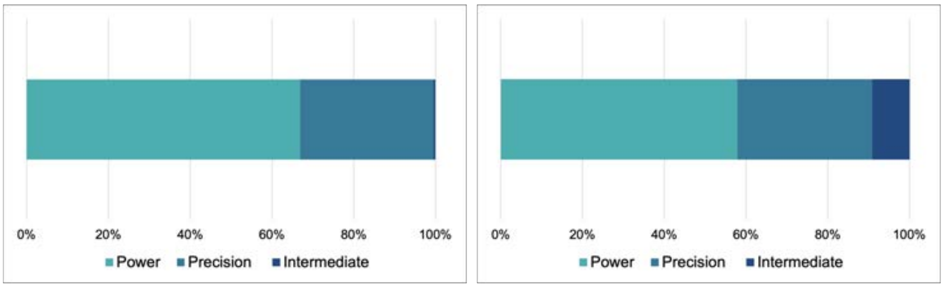
- Parts: Statistical significance was found between object categories: χ^2 **(2, N = 1872) = 77.17, p < 0.001*** with One-Part being predominantly grasped using Power grasps (66.88%, N = 626), followed by Precision grasps (32.58%, N = 305) and Intermediate grasps (0.53%, N = 5) and Multi-Part being predominantly grasped using Power grasps (57.9%, N = 529), followed by Precision grasps (33.01%, N = 309) and Intermediate (9.08%, N = 85).

Figures 8.8, 8.9 and 8.10 show the split between Power, Precision and Intermediate grasps for each object category presented in this experiment.



(a) Stable (b) Unstable

Figure 8.9: The use of Power, Precision and Intermediate grasps for VOE object categories: Stable and Unstable



(a) One-Part (b) Multiple-Part

Figure 8.10: The use of Power, Precision and Intermediate grasps for Parts object categories: One-Part and Multiple-Part

8.9.2.2 Most Used Grasp Types

Figure 8.11 shows the most common grasp types and their usage percentages. Large Diameter [P1] was the most prevalent grasp type, accounting for 28.09% of the labels in the data set (N = 526). This grasp was followed by the Small Diameter [P2] grasp with 11.11% of the labels (N = 208) and the Medium Wrap [P3] from the Power grasp category with 10.30% of the labels (N = 193). The five most used grasps accounted for 67.30% of the data. The remaining instances were grasped using Thumb-Finger Variation grasps (16.77%), Sphere Variation grasps (10.52%) and Other (5.39%), as defined in the VR Taxonomy of Grasp Types presented in Chapter 6.

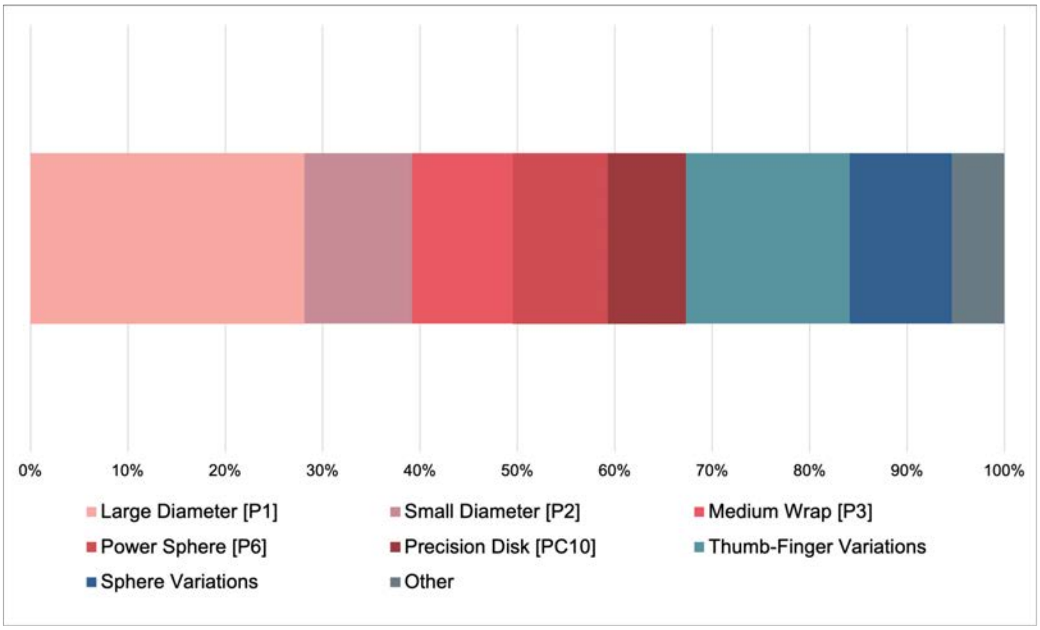


Figure 8.11: Most used grasp types in this experiment showing the frequency of the five most used grasp types in the experiment: Large Diameter [P1], Small Diameter [P2], Medium Wrap [P3], Power Sphere [P6], Precision Disk [PC10] and the frequency of grasp categories defined in the first VR Taxonomy of Grasp Types: Thumb-Finger Variations, Sphere Variations and Other.

8.9.2.3 Analysis - Grasp Labels

Considerable effort has been made in determining how certain parameters influence human grasping when manipulating real objects (Feix et al., 2014b, 2014a) and it was found that grasp choice is at least 43% influenced by object properties and 31% influenced by the task properties (Feix et al., 2014a). However, results in this work show that grasp choice was not influenced by different docking conditions, being primarily influenced by object characteristics, with statistically significant differences being found between object categories for all categorisation methods under experiment. For Zingg's methodology (Zingg, 1935), results are consistent with results in Chapter 6, where Oblate and Equant objects presented a similar grasp category segmentation, with around 70% of grasp instances being from the Power grasp category, while the remaining choices were from Precision category. Unique patterns were found for Bladed category where the split between Precision and Power being $\approx 50\%$ and Prolate category where grasps from all grasp categories (Power, Precision and Intermediate) were used.

For VOE categorisation method, Stable was predominantly grasped using Power grasps, followed by Precision, however Unstable showed a unique pattern, being grasped with Power, Precision and Intermediate grasps, with the use of Intermediate grasps increasing to $\approx 20\%$. This indicates a higher variability in grasp category for objects that are curved around certain axes. A similar pattern was found for Parts categorisation method, where Multiple-Part showed higher variability in grasp category, with objects in this category being grasped using Power, Precision and Intermediate grasp types, which is consistent with real grasping literature showing that complex objects (or irregular shapes) present high variability in grasp type chosen (Feix et al., 2014b; Kamakura et al., 1980).

Results presented in this chapter show that five grasp types accounted for $\approx 70\%$

of the data set. This finding is consistent with results presented in Chapter 6 and shows a lower variability of grasps when manipulating objects in VR compared to reality, where grasp studies showed that 13 grasp types are needed to account for 80% of the data (Feix et al., 2014b). When interacting with real objects, sensory feedback (shape and mass) constrains the grasping choice, however, the lack of haptic feedback in VR (Cooper et al., 2018) allows more freedom in terms of grasping choice (for example, a large box could be grasped using a tip pinch grasp [PC11], without any consequences caused by the mass or shape of the object as the hand is allowed to penetrate the virtual object any time during the interaction). The lack of these constraints may be the reason for grasping patterns showing less variability in VR compared to real environments. This lower variability in grasp approach was also found in grasping virtual objects in MR (Al-Kalbani et al., 2016a) and in gesture elicitation studies (Billinghurst et al., 2014), where subjects used a small variety of hand poses to complete various tasks.

8.9.3 Taxonomy of Grasp Types

The first VR Taxonomy of Grasp Types presented in Chapter 6 was structured following the methodology for developing real grasping taxonomies presented in (Feix et al., 2009, 2014a, 2014b), taking into consideration findings from the user experiment and grouping the grasp types in representative categories: Main VR grasps, Thumb-Finger variations and Sphere variations. Findings in this experiment showed similar patterns in the frequency of grasps used, therefore the main categories of the taxonomy are unchanged, with grasp types in each category being updated to reflect findings in this experiment. The Main VR grasps present the five most common grasp types used in this experiment: Large Diameter [P1], Small Diameter [P2], Medium Wrap [P3], Power Sphere [P6] and Precision Disk [PC10]. Thumb-Finger Variations category shows the same grasps identified in Chapter 6 with the Stick [I5] grasp being moved from Main VR Grasps in this category due to being less frequently used: Thumb 2-Finger [PC4], Stick [I5], Thumb 3-Finger [PC5], Thumb 4-Finger [PC6], Thumb-Index Finger [PC1] and Tip Pinch [PC11]. In this experiment, subjects used more Sphere Variations grasps, therefore this category was updated to contain the following: Tripod [PC7], Power Disk [P5], Quadpod [PC8], Precision Sphere [PC9], Inferior Pincer [PC2] and Sphere-4 Finger [P7]. Moreover, the taxonomy presented in this chapter shows grasping patterns based on object categorisation methods (Zingg, VOEquilibrium, Parts). Figure 8.12 presents the updated VR Taxonomy of Grasp Types which shows correlations between object category and grasp type with percentages and number of instances for each object category.

Method	Grasp Type / Object Category	Main VR Grasps [67,32%]					Thumb-Finger Variations [21,31%]						Sphere Variations [10,52%]						Other [0,85%]
		[P1]	[P2]	[P3]	[P6]	[PC10]	[PC4]	[I5]	[PC5]	[PC6]	[PC1]	[PC11]	[PC7]	[P5]	[PC8]	[PC9]	[PC2]	[P7]	[P13] [I4]
Zingg	Equant	25,43% [119]	8,97% [42]	2,99% [14]	32,69% [153]	1,28% [6]	2,35% [11]	0,00% [0]	2,35% [11]	2,99% [14]	1,92% [9]	0,64% [3]	3,63% [17]	0,64% [3]	3,85% [18]	6,84% [32]	1,71% [8]	1,07% [5]	0,64% [3]
	Prolate	0,21% [1]	25,43% [119]	28,63% [134]	1,50% [7]	0,21% [1]	7,91% [37]	18,16% [85]	4,91% [23]	4,06% [19]	3,21% [15]	1,07% [5]	1,71% [8]	0,00% [0]	0,21% [1]	0,00% [0]	0,21% [1]	0,00% [0]	2,56% [12]
	Oblate	45,51% [213]	4,49% [21]	7,69% [36]	2,78% [13]	4,70% [22]	5,34% [25]	0,00% [0]	7,26% [34]	4,70% [22]	2,99% [14]	1,07% [5]	4,70% [22]	4,27% [20]	2,78% [13]	0,64% [3]	0,85% [4]	0,00% [0]	0,21% [1]
	Bladed	41,24% [193]	5,34% [25]	2,14% [10]	2,14% [10]	25,85% [121]	5,77% [27]	0,00% [0]	2,99% [14]	2,99% [14]	1,28% [6]	1,28% [6]	2,56% [12]	4,06% [19]	1,28% [6]	0,21% [1]	0,85% [4]	0,00% [0]	0,00% [0]
VOE	Stable	37,39% [525]	8,55% [120]	8,12% [114]	5,98% [84]	10,61% [149]	5,27% [74]	0,07% [1]	4,56% [64]	3,70% [52]	2,35% [33]	1,00% [14]	3,85% [54]	2,92% [41]	2,49% [35]	1,50% [21]	1,14% [16]	0,21% [3]	0,28% [4]
	Unstable	0,21% [1]	18,59% [87]	17,09% [80]	21,15% [99]	0,21% [1]	5,56% [26]	17,95% [84]	3,85% [18]	3,63% [17]	2,35% [11]	1,07% [5]	1,07% [5]	0,21% [1]	0,64% [3]	3,21% [15]	0,21% [1]	0,43% [2]	2,56% [12]
Parts	One-Part	29,81% [279]	4,91% [46]	13,25% [124]	16,88% [158]	7,59% [71]	4,06% [38]	0,11% [1]	5,34% [50]	3,74% [35]	1,28% [12]	0,64% [6]	2,88% [27]	1,50% [14]	2,56% [24]	3,53% [33]	0,96% [9]	0,53% [5]	0,43% [4]
	Multiple-Part	26,39% [247]	17,20% [161]	7,48% [70]	2,67% [25]	8,44% [79]	6,62% [62]	8,97% [84]	3,42% [32]	3,63% [34]	3,42% [32]	1,39% [13]	3,42% [32]	2,99% [28]	1,50% [14]	0,32% [3]	0,85% [8]	0,00% [0]	1,28% [12]

Figure 8.12: VR Taxonomy of Grasp Types. Grasps are categorised by frequency, showing percentage and number of instances for each object category from the three categorisation methods presented: Zingg, VOEquilibrium and Parts.

8.9.3.1 Analysis - Taxonomy of Grasp Types

Following the user elicitation and labelling methodologies presented in Chapter 4, an updated VR Taxonomy of Grasp Types is proposed to encapsulate different object characteristics. The taxonomy was structured following methodologies for defining taxonomy dimensions in real grasping literature (Feix et al., 2009) and are consistent with dimensions proposed in the first VR Taxonomy of Grasp Types presented in Chapter 6. While real grasping research showed that both task (Feix et al., 2014a) and object characteristics (Feix et al., 2014b) influence grasping patterns, work presented in this thesis showed that virtual grasping patterns are primarily influenced by object characteristics, therefore the taxonomy presented in this chapter presents grasping patterns for virtual objects based on their shape, stability and complexity. Figure 8.12 presents the final VR Taxonomy of Grasp Types which encapsulates these three categorisation methods and provides an overview of the grasp types used in VR, with only 17 grasp types being used for virtual objects out of the 34 grasp types available in the most complete human grasp taxonomy to date (Feix et al., 2009).

Since results in this experiment showed no significant statistical differences between docking tasks, the main VR grasps identified in this chapter are consistent with the main grasps identified in Chapter 6. The main difference in this updated taxonomy is that Thumb-2 Finger [PC4] was more frequently used than Stick [I5], which was therefore moved to the Thumb-Finger Variations category since the hand posture in Stick [I5] is another variation of positioning the fingertips around the object. As discussed in Chapter 6, subjects used variations of the same grasp but using different number of digits to perform precision and intermediate grasps, which is in alignment to interaction findings in augmented reality where users did not show awareness of the number of digits used during interactions

(Piumsomboon et al., 2013). The Large Diameter [P1] remains the most used grasp type in VR, as opposed to real environments where the Medium Wrap [P3] is the most common grasp (Feix et al., 2014b).

When analysing how grasp pattern changes between object characteristics categories, unique patterns are identified for different groups. For Zingg's categorisation method (Zingg, 1935), differences were identified in terms of grasp labels between object shape categories, with patterns being consistent with results presented in the first VR Taxonomy of Grasp Types in Chapter 6. For VOEquilibrium method, the most used grasp type for the Stable category is Large Diameter [P1], while the most used grasp type for Unstable is Power Sphere [P6]. This might have been influenced by the shape of Unstable objects, being spherical/cylindrical and therefore influencing users to curve their hand in a position suitable for holding a round or sphere-shaped object. This is consistent with real object grasping literature, where Power Sphere [P6] grasps are predominantly used for spherical objects while Large Diameter [P1] grasps are predominantly used for cuboid/prismatic objects (Cai, Kitani, & Sato, 2016).

Parts categorisation method also showed significant differences in grasp labels between object categories; One-Part objects were predominantly grasped using Large Diameter [P1] and Power Sphere [P6] and Multi-Part objects were predominantly grasped using a Large Diameter [P1] and Small Diamter [P2] showing more variability in the use of Thumb-finger variations. This is consistent with assumptions made in Chapter 6 and 7, claiming that higher variability in objects such as Mug, Clamp, Scissors is due to increased object complexity and the presence of handles and protruding objects. This is also consistent with real object grasping where objects with more complex shapes present higher variability in grasp types (Feix et al., 2014b). This effect might be motivated by Biederman's theory (Biederman, 1987) which shows that humans' visual system naturally decomposes structured objects into component parts. Therefore, it can be assumed that the change in grasp pattern between One-Part and Multiple-Part is due to users visually decomposing the objects in smaller components and adjusting their grasp based on the shape and size of that component.

8.9.3.2 Post-test Questionnaire

When asked what properties influenced their approach to grasp the virtual objects, *interaction object characteristics* was top rated, followed by *target object rotation and position*. Further, when asked to describe their strategy for grasping, some participants reported: *I picked up objects from a specific angle so that I'd be able to rotate my hand to place them in the correct rotation. To do this I checked where the objects had to go before picking them up.* [P14]; *The VR experience felt very much like the real world and the positions of the targets completely influenced my decision regarding the way I grasp the object* [P27]. The responses to the strategy of grasping can be found in full in Appendix C.

8.10 Discussion and Conclusions

This chapter analysed grasping patterns for object characteristics such as shape, stability and complexity on grasping patterns. Additionally, since real grasping literature showed that both object characteristics and task influence grasping patterns (Feix et al., 2014b, 2014a), this work explored grasp metrics for docking tasks in VR, which has shown to be the preferred interaction task in VR due to following a direct mapping to real world tasks (Suhail, Sargunam, Han, & Ragan, 2017). Results showed that grasping patterns did not change for different rotation and translation docking tasks, however, grasp metrics showed to change with object characteristics.

When investigating grasping real objects, researchers often used circular and bar objects (Bullock et al., 2015) or a sample of daily objects (Matheus & Dollar, 2010) and report representative grasp patterns for each object type (Feix et al., 2014b). However, while real environments show the same level of fidelity for different scenarios and objects, in VR there can be different platforms, tracking sensors, rendering processes and visualisations, which provide different levels of fidelity and realism. Therefore, developing a generalizable model for object categorisation which can lead to an improved natural grasping interface could support a greater quality of experience across devices/platforms. To achieve this, the VR Taxonomy of Grasp Types presented in this chapter provides an overview of grasp choice for different object categories.

Similar to real object grasping (Feix et al., 2014b, 2014a), virtual objects with similar characteristics showed to have a specific grasping pattern associated with them; When analysing grasping patterns for object shape, results showed that Equant objects are generally grasped using a Power Sphere [P6] and a Large Diameter [P1] being predominantly grasped on dimension C; Prolate objects are generally grasped using a Medium Wrap [P3] being predominantly grasped on dimension C and B; Oblate objects are generally grasped using a Large Diameter [P1] showing higher variability in the use of Thumb and Sphere variations being predominantly grasped on dimension C; Bladed objects showed a unique pattern, being predominantly grasped on dimension B with a Large Diameter [P1] and Precision Disk [PC10] which is consistent with real grasping behaviour where disk-shaped objects are predominantly grasped on B with a Disk variation grasp ([PC10] or [P5]) (Kamakura et al., 1980).

When analysing grasping patterns for object stability, Stable objects were predominantly grasped using a Large Diameter [P1] on dimension C, while Unstable objects being grasped using a Power Sphere [P6] on dimension C. This shows that users mimicked real grasping behaviour and adapted their grasp to the cylindrical/spherical shape of the virtual object. When analysing object complexity, both object categories showed a prevalence for Large Diameter [P1] and dimension C, with Multi-Part objects showing higher variability in grasp types chosen and a more balanced split between grasps around dimension C and B, which was unsurprising since irregular and more complex shapes show higher variability in real grasping due to multiple graspable parts available (Feix et al., 2014b). However, when categorising objects into Parts categories a higher variation between users was observed (as shown in Table 8.1) which might illustrate further complexity into automatically categorising objects using Parts categorisation method.

Considering the above, the null hypothesis H_2**: Grasping patterns in VR are not influenced by rotation and translation docking tasks** is rejected and the hypothesis H_1**: Categorisation based on object characteristics influences grasping patterns in VR** is accepted due to both grasped dimension and grasp labels being influenced by object categories. These differences were synthesized in an updated VR Taxonomy of Grasp Types developed iteratively based on existing methodologies for collecting and structuring data in a taxonomy, detailed in Chapter 4. Moreover, this taxonomy provides insights into how categorisation methodologies are an important tool in supporting the development of automatic object classification and could support automated or semi-automated adaptation and improvement of VR grasping methods.

9 | Conclusions and Recommendations

The primary aim of this work was to evaluate grasping patterns in VR and develop the first VR Taxonomy of Grasp Types. This was achieved through reviewing existing methodologies for analysing grasping patterns in real environments and existing taxonomies in HCI and using this knowledge to define a methodology for developing a VR grasp taxonomy that takes into consideration VR constraints (Chapter 4).

Following this methodology, grasping patterns in VR were compared to real grasping patterns to understand the fundamental similarities and differences between the two (Chapter 5). Findings from this work showed that there are similarities such as users adapting their grasp patterns to object characteristics, as well as differences such as users grasping with more power grasps in VR and more precision grasps in real environments, as well as errors in estimating object size due to the lack of haptic feedback. To fully understand how users adapt their grasp based on virtual object characteristics, grasping patterns for object shapes were analysed and organised in the first VR Taxonomy of Grasp Types (Chapter 6). Apart from significant differences in grasp metrics between object categories, findings in this work showed that subjects presented higher variability in grasping for objects that present multiple graspable locations, such as protruding objects or handles. Moreover, the lack of haptic feedback showed to influence grasping patterns in VR, users over and underestimating object size.

Visual cues for thermal haptic feedback were explored to identify changes in grasping patterns for these conditions, which are particularly important for training and simulations (Chapter 7). Results in this work showed that thermal visual cues influenced the grasp metrics under analysis, which might change the original taxonomy when thermal cues are presented on the virtual object. To further improve the first taxonomy presented, another study was conducted to explore how grasping metrics change for virtual object categorisation methods in a mixed docking task (Chapter 8). Findings showed that different object categories for shape, stability and complexity present different grasping patterns, which were synthesized in a more complete VR Taxonomy of Grasp Types.

This chapter will discuss recommendations and implications of this work into the HCI and VR community, as well as highlighting limitations and future work. This chapter is structured as follows: Section 9.1 presents recommendations for interaction design based on findings in this thesis, Section 9.2 presents a summary and Section 9.3 presents limitations and future work.

9.1 Recommendations

To complement the findings discussed across the four user studies in this thesis (Chapter 5-8), this section presents recommendations and transferable implications of this work on the VR research community and future VR grasping models. These are based on nine key observations presented in this section.

Grasp aperture changes with different virtual objects, the main influencing factor being size, with subjects under and overestimating virtual object size when grasping in VR. Chapter 5 showed that there is a high variability in GAp for all virtual objects used in the experiment, with unique patterns being identified for specific objects. This shows that when choosing a GAp users took object characteristics into consideration and adapted their hand opening accordingly, which is consistent with real grasping literature (Feix et al., 2014b) and freehand gestures research where participants increased the size of their gesture when the size of the object increased (Pham, Vermeulen, Tang, & MacDonald Vermeulen, 2018).

Yet, when analysing intuitive grasping, Al-Kalbani et al. (Al-Kalbani et al., 2016a) found that subjects used a common GAp regardless of virtual object size. However the authors used abstract objects such as spheres and cubes for interaction, while the work in this thesis provided users with virtual representations of daily objects. This adaptation of GAp based on virtual objects observed in this thesis can be attributed to the fact that users are more likely to interact with more accuracy with familiar objects than unfamiliar ones, as familiar objects can enable users to use the size and shape of familiar objects as a cue for size and distance estimation (II, Kuparinen, Rapson, & Sandor, 2017; H. Park et al., 2021).

While they adapted their grasp for various object sizes, subjects over and underestimated object size in most of grasp instances, creating grasp apertures both larger and smaller than object dimensions. The finding of creating smaller grasp apertures is in agreement with VR literature where users often underestimate distance and size in VR (Murgia & Sharkey, 2009; Masnadi, Pfeil, Sera-Josef, & LaViola, 2021), however, the larger grasp apertures might be attributed to behaviour learned from realistic scenarios (Jacob et al., 2008). In reality, we approach interactions with caution, waiting for haptic feedback to guide us (e.g. we do not overshoot when grasping a mug, instead we create a larger aperture and use vision and haptics to adapt it into a secure grasp), however, where haptic feedback is missing in VR, there is no real indication on where the grasp should end, therefore leading to larger grasps when visual feedback is not entirely reliable due to occlusion.

To understand if there is a pattern in over and underestimating object size, Chapter 6 presented an analysis into GAp for various object shapes. Findings showed that there was no clear pattern based on object shape, however it proposed the question of whether this variability is linked to the location of the grasp (e.g objects that have handles would more likely be grasped larger and smaller due to users grasping on different locations). However, in Chapter 7 where grasp location was explored, the same effect was observed: users grasp virtual objects larger and smaller than object dimensions regardless of the location of the grasp. Yet, it has been observed that smaller objects are grasped larger than object dimensions, which was found in all experiments presented in this thesis. However, future work should explore if this applies to all small objects and understand at what object sizes subjects start to also grasp objects smaller.

Findings in this work indicate that interaction designers should be aware of the potential discrepancies in virtual size estimation during grasping. These discrepancies using GAp present a fundamental problem for grasping interaction, and can potentially be attributed to the size distortion that is caused by the virtual object perception. Grasping approaches that check for collisions between the hand and the virtual object (Höll, Oberweger, Arth, & Lepetit, 2018; Furmanek et al., 2021) should consider triggering interaction inside and outside the bounds of the virtual object, to mitigate estimation errors when haptic feedback is missing.

Users commonly replicated realistic grasp approaches rather than biasing specific less natural grasps of the virtual object. Findings in user experiments presented in this thesis showed that users generally tried to mimic real grasping behaviour as opposed to performing gestures such as pinch, which are commonly used in current virtual environments to detect intended grasp intentions (Moehring & Froehlich, 2011). One popular example of such gestural interface is the Microsoft Hololens (Avila & Bailey, 2016; Yim et al., 2016). However, in this work, users adapted their grasps based on object characteristics such as shape, stability, object complexity, as well as grasping the objects on various locations, and by wrapping their hand around the object, with the most common grasp type being

Large Diameter (where the fingertips are far from the thumb, positioned to hold a large cylindrical shape). This is consistent with freehand gesture work, where researchers found that users often performed gestures that are representative of real interactions (Pham et al., 2018), which could be attributed to the ideas of Reality-Based Interaction (RBI) (Jacob et al., 2008) where realistic virtual environments influence users to behave similarly to real life interactions: for example, in the work of Pham et al. (Pham et al., 2018) authors observed that since in reality it does not make sense to reach through or under a physical object on a table, subjects in their elicitation study avoided these types of interactions with virtual objects. This can also be linked to GAp findings where subjects grasped smaller objects larger in all grasp instances, which might be due to users avoiding abstract gestures (such as pinch where GAp would be near 0) and instead wrapping their hand around the object.

While RBI theory shows that new interaction styles draw strength by building on users' pre-existing knowledge of the every-day, non-digital world (Jacob et al., 2008), the extent to which users will replicate real-life knowledge in VR is influenced by the nature of referents used for interaction. For example, (Piumsomboon et al., 2013) showed a prevalence for abstract and distance gestures in an elicitation study where some of the referents used were not objects from daily life such as menus and icons. It is important therefore to note that findings in this thesis might not be applicable for interactions with abstract objects, which do not have any learned patterns from real environments (for example, in virtual environments where interactions might be required with non-solid or non-rigid objects i.e gaseous objects, fluid objects).

Moreover, these findings can also be attributed to the methodology used in the elicitation studies, where users were asked to grasp the virtual objects in the most intuitive way, therefore it can be assumed that in a completely free grasping VR scenario, users might perform different interactions with virtual objects. However, the consistency across all studies and tasks in this thesis, illustrates that the patterns uncovered may be more generalisable. Taking these into consideration, interaction designers should use these findings for virtual environments where grasping interaction is needed to achieve high immersion and presence levels and reality is closely replicated, such as for training (Nayer et al., 2020) and simulations (Barkokebas, Ritter, Sirbu, Li, & Al-Hussein, 2019).

Virtual object characteristics (shape, stability and complexity) appear as the primary influencing factor for grasping patterns in VR. Chapter 5 showed that users adapted their grasp approach based on virtual object shape and size, showing similarities to real grasping approach, which is consistent with RBI theory (Jacob et al., 2008) where subjects tend to mimic behaviours from real life in realistic VR environments. Chapter 6 then explored how object shape influences this approach and found different patterns for different object shapes during simple translate tasks. To evaluate if docking tasks influence grasping patterns, Chapter 8 explored grasp metrics during a translate and rotate task, and found that task did not influence grasping in VR, with results being similar to the first iteration of the grasp taxonomy presented in Chapter 6. This is contrary to real grasping literature where task is one of the primary influencing factor in grasp metrics (Napier, 1956; Feix et al., 2014a).

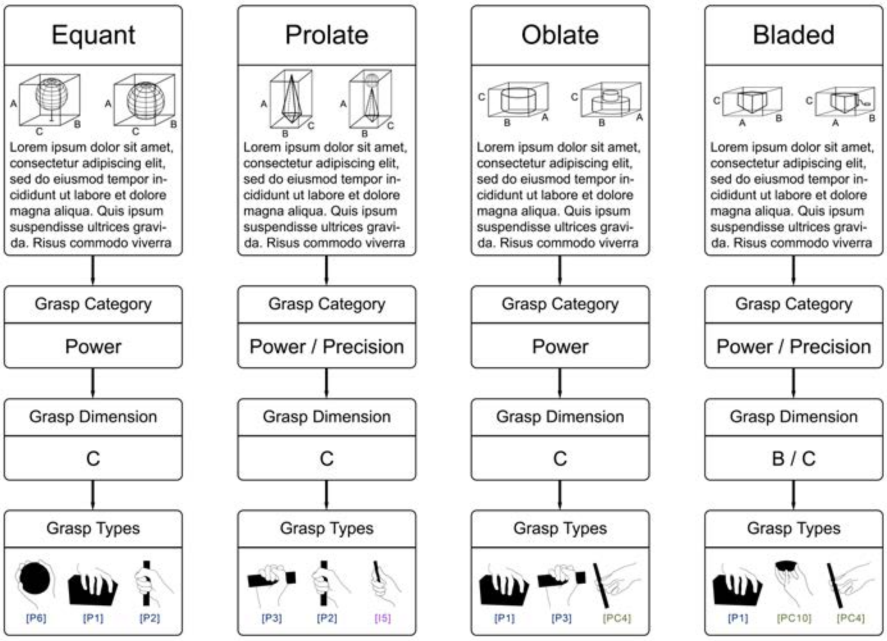
Chapter 8 also explored additional object categorisation methods to further evaluate the influence of virtual object characteristics on grasping metrics in VR. Results showed that users chose different patterns for different categories of objects, revealing significant differences in grasping approach for all object categories explored in this study. Following the work in Chapter 7 which revealed that users grasp differently based on the location of the grasp, Chapter 8 also explored grasped dimension and how this changes between different object categories.

These results were synthesized and presented in Figure 9.1, which could serve as a framework of grasping interaction for 3D interaction designers. Using this knowledge, researchers could quickly identify the most intuitive grasp category and type for objects that fall in one of the categories explored in this work and therefore develop hand poses that fulfil the requirements of these grasps for natural and intuitive virtual object manipulation in VR.

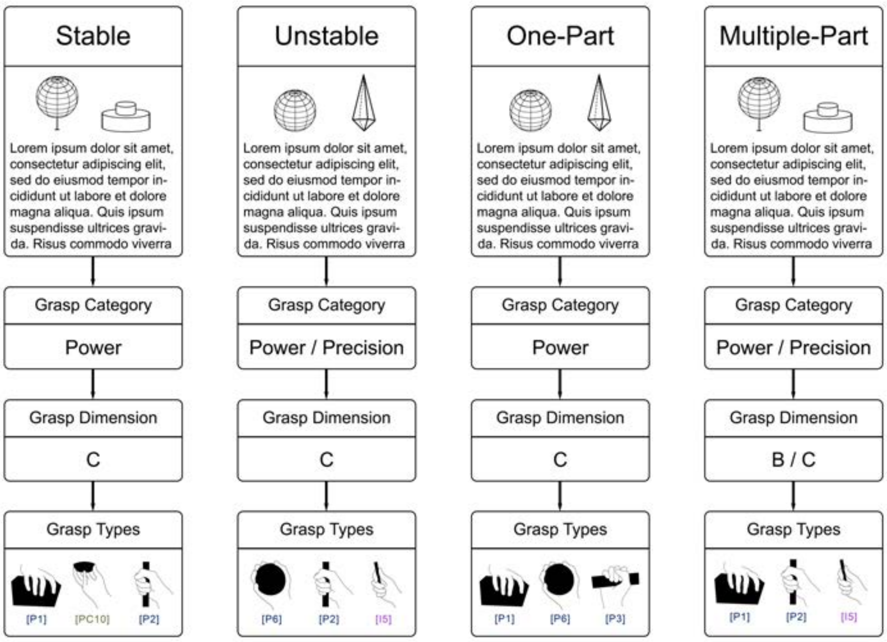
Additionally, this knowledge can also guide the design and construction of virtual environments and objects. For example, in environments that have only Equant objects, the three main grasps shown in Figure 9.1 should be expected for interaction, or for virtual environments where multiple types of objects are presented, systems that adapt them to fit in one category and therefore be grasped with a limited number of grasps might be implemented to provide a more simplistic interaction.

Moreover, grasp dimension can be used for calculating the optimal point of contact between the fingers and the virtual object. Collision detection algorithms have shown to be complex and computationally expensive, with current physics engines still being limited in supporting collisions for concave mesh colliders (Höll et al., 2018). However, using the knowledge in this thesis, interaction designers could focus on developing abstract mesh colliders that cover only the graspable dimension of the virtual object (e.g. dimension C for equant objects).

Object categorisation methods can inform interaction design decisions when creating intuitive grasping interactions in VR. The methods proposed for categorisation in this thesis focus on attributes that are relevant to grasping, and therefore could be used by designers creating VR environments that require di-



(a) Zingg's object categories



(b) VOEquilibrium and Parts object categories

Figure 9.1: Grasp patterns categorisations and most found trends.

rect interaction with objects to categorise the objects required in the environment based on these attributes. These objects should then be linked to the corresponding grasp types of each shape category presented in the experiments of this thesis, rather than a generalised grasp model which might lack naturalness and be computationally complex (Tian et al., 2019).

It is known that a high degree of hand pose recognition is achievable, but computationally expensive (Piumsomboon et al., 2013), therefore, computer systems aiming to achieve seamless and natural interactions in VR could benefit from categorising objects and focusing on the most prevalent grasp poses for each category presented in the first VR Taxonomy of Grasp Types detailed in this thesis. Apart from grasp poses, the categorisation methods could also be used for mesh and collider design and development. When developing real-time grasping interactions in VR, the data structure of the colliders should occupy as little memory as possible (Ericson, 2004), which makes the process of developing a generalisable collision detection algorithm that is computationally unexpensive quite challenging. Considering findings in this work, researchers could categorise virtual objects based on their attributes and develop appropriate colliders for each category, to meet the requirements of the expected grasp pose and location.

Since machine learning algorithms have proven to be useful for object classification (Farid & Sammut, 2013; Sahbi, Audibert, & Keriven, 2011), these categorisation methods could be furthered into a machine learning algorithm to automate this correlation of grasp types to object attributes, to allow natural and intuitive interactions in VR.

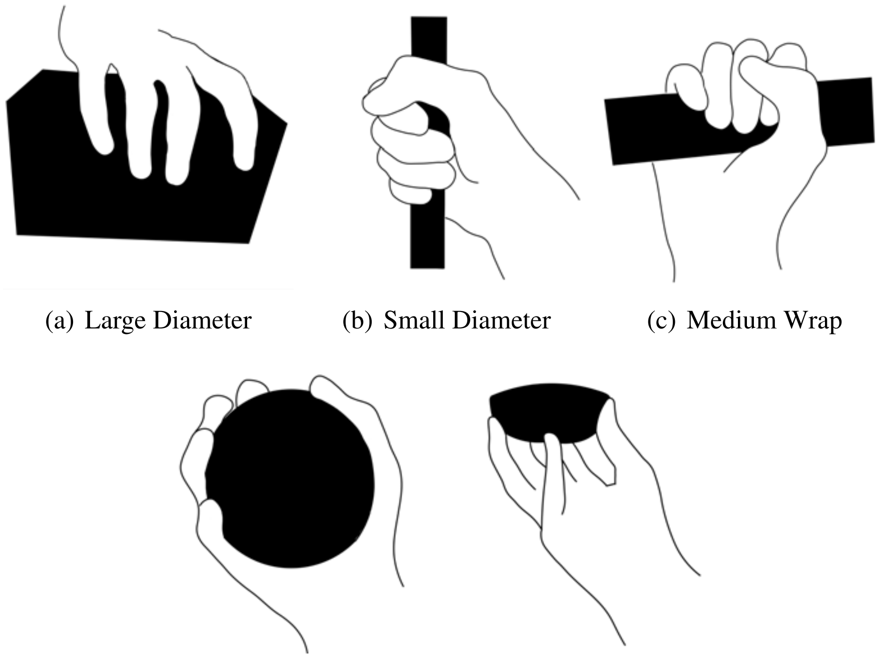
The most common grasps identified for virtual grasping are power grasps. Chapter 5 showed that real objects are predominantly grasped using precision grasps while the same objects represented in 3D in VR are predominantly grasped with power grasps. This prevalence for power grasps was maintained in all the elicitation studies presented in this thesis, showing that regardless of virtual object characteristics, subjects are more likely to use power grasps in VR. This is in agreement with freehand grasping literature in immersive environments where users generally grasp virtual objects with a power grasp (Al-Kalbani et al., 2016a). The lack of sensory feedback in VR might have influenced this result, as there is no need for applying any force or precision to perform a stable grasp if the object has no perceived weight or texture (Cooper et al., 2018). Moreover, when there is no haptic feedback at the end of the grasp, users do not need to focus on the fingertips to feel when the grasp is stable and instead they can perform a more relaxed hand pose where the fingertips are more relaxed.

Power grasps are linked to stability and security, being distinguished by large areas of contact between the hand and the object (M. R. Cutkosky, 1989) while precision grasps present grasps where the object is commonly held between the fingertips to allow an increased level of manipulation. While the work presented in this thesis provides much more detail on hand poses for grasping interactions in VR, 3D interaction designers could use this insight when a generalised grasp model is required. For example, in environments where achieving natural grasping interactions is not cost or computationally effective, such as for consumer and entertainment VR (Hock, Benedikter, Gugenheimer, & Rukzio, 2017), designers should consider user preference for power grasps and enable hand poses where the palm comes in contact with the virtual objects and the fingers wrap around it as in a claw (which would be a variation of a Large Diameter, the most used grasp type in this work) as opposed to focusing on interactions between fingertips and object bounds.

Only five grasps account for the majority of grasp data in VR. Chapter 6 showed that only six grasps account for more than 80% of the data when performing simple translate tasks in VR. Five of these grasps were found again as the main grasp types which accounted for more than 65% of the grasp data when performing docking tasks in VR. These grasps are Large Diameter, Small Diameter, Medium Wrap, Power Sphere and Precision Disk (see Figure 9.2) and were the most common grasp types across all VR user experiments presented in this thesis. When analysing studies involving grasping objects in real life, a total of 13 different grasp types accounted for $\approx$ 80% of the data (Feix et al., 2014b), with a total of 34 grasp types being reported in the most complete grasp taxonomy for real objects (Feix et al., 2009), suggesting that grasp choice in VR varies less than in real environments.

It has been shown that having a generalised grasp model can hinder usability and accessibility, researchers showing that when interacting with virtual objects, even when instructed to perform a specific grasp, user's grasp approach varies considerably (Al-Kalbani et al., 2016a). This was evident in the work presented in this thesis as users adapted their grasp based on influencing factors such as object characteristics and thermal cues and therefore suggests that interaction designers should consider allowing more than one grasp for intuitive environments. While this can be computationally expensive, computer systems aiming to achieve seamless and natural interactions in VR could benefit from these results by only focusing on the five grasps when designing grasping interactions in VR.

Results indicate that users did not show an awareness of the number of fingers used when grasping in VR. While results in this work showed less variability in the number of grasp types used for virtual objects, it can be observed that variants of a single hand pose were often used across multiple objects and



(a) Large Diameter (b) Small Diameter (c) Medium Wrap

(d) Power Sphere (e) Precision Disk

Figure 9.2: The 5 most common grasp types used for intuitive interaction in VR.

participants, especially when looking at the number of fingers used to perform a precision grasp. This is consistent with elicitation studies for touch gestures (Wobbrock et al., 2009) and augmented reality (Piumsomboon et al., 2013) where participants did not show an awareness of the number of fingers used while interacting with virtual content. This contrasts with grasping and manipulating real objects, where the number of digits involved is influenced by the size of the object (Bullock et al., 2015), increasing with size and mass (Cesari & Newell, 1999, 2000).

This finding had an influence on defining the dimensions of the taxonomy, with the second most important category of grasps in VR being Thumb-Finger Variations. This is an important contribution for 3D grasp interaction design, especially for VR environments where natural and intuitive interaction highly influences the outcomes of the VR experience such as training and simulations. In these environments, for interactions where precision grasps are required, designers could focus on triggering interaction by only focusing on the thumb and index finger. If these two fingers are engaged in the grasp (where engagement should also consider grasp dimension and aperture for calculating the grasp point), then whether or not the other fingertips engage in the interaction should be irrelevant from a grasp trigger point of view. These assumptions should however be tested against a benchmark grasp model to understand the usability issues that might arise.

Visual cues for thermal representation influence users to grasp objects differently in VR. Chapter 7 showed differences in grasp metrics when visual cues for temperature were displayed together with a virtual mug. Subjects showed to grasp the virtual mug smaller when thermal cues for hot were displayed, as compared to empty and cold conditions. This finding was linked to users grasping the handle more often with visual cues for hot than with other conditions, which explained the smaller grasp aperture. This change in the grasp location also influenced the grasp types used for each condition, users grasping the handle with more precision grasps and the body and top of the mug with more power grasps.

While differences were observed between thermal conditions for all grasp metrics, in fact, thermal cues only influenced the grasped location, users mimicking real grasping behaviour and therefore predominantly grasping the handle of the mug to avoid getting burnt (Jacob et al., 2008). Then, the grasped location influenced grasp aperture and grasp types, users adjusting their grasp to the size and shape of the grasped object (handle, top or body of the object), which is consistent with findings from Chapter 6 and 8 where it has been shown that object characteristics influence grasp metrics in VR. This finding is important for designing interactions in VR training or simulations, especially as they are increasingly used for health and safety applications such as training for hazardous situations (Shaw et al., 2019), especially for environments where the user needs to manipulate thermally variable objects (Bharath & Patil, 2018; Price, Kuttolamadom, & Obeidat, 2019). Interaction designers in these environments could trigger interactions on different object locations, taking into consideration that users might avoid thermally dangerous objects in realistic VR scenarios.

This thesis presented the first VR Taxonomy of Grasp Types which contributes to achieving natural and intuitive VR interactions. Following current methods for analysing grasp metrics and developing taxonomies in HCI, this work proposed a method for developing grasp taxonomies for VR by conducting elicitation studies where users are asked to grasp virtual objects in an intuitive way, under different conditions. Then, the grasps collected are labelled using grasp metrics from state-of-the-art grasp taxonomies for real objects. This work explored how subjects intuitively grasp virtual objects based on object characteristics, translate and docking tasks and visual thermal cues, however, this methodology could be further used to build on the existing virtual grasping knowledge and understand how users intuitively grasp objects under various conditions, relevant for the VR system in question.

This work first proposed a grasp taxonomy based on object shape (Zingg, 1935) and for a simple translate task (Chapter 6), which was furthered in a more complete grasp taxonomy based on object shape, stability and complexity and for translation and rotation tasks (Chapter 8), presenting the first VR Taxonomy of Grasp Types. This taxonomy aims at taking a step forward in filling a research gap, as no previous work explored how users intuitively grasp virtual objects, to inform design decisions when developing intuitive grasp models in VR. This taxonomy might benefit researchers in providing insights into the main hand poses and object locations which should trigger grasping interaction in VR, therefore providing a guideline for understanding virtual grasping patterns in more detail. Moreover, this taxonomy can be used as a framework for comparative analysis of freehand grasping-based interaction techniques, providing researchers with the main categories of grasps that should be further explored in VR: the five main VR grasps, generally used for power, relaxed hand poses and thumb-finger variations used for precise, fine manipulations in VR.

9.2 Summary

This thesis presented the first analysis of grasping patterns in VR, with results being synthesized in the first VR Taxonomy of Grasp Types. The first main contribution of this thesis is providing a method for collecting grasp data in VR, classifying the grasps collected based on current grasp metrics used for analysing grasping patterns in reality and synthesizing the results in a taxonomy. This method was used across four user experiments to explore changes in patterns based on various parameters and can be further used for developing taxonomies for different VR configurations, other interaction tools such as wearable devices or future improved devices to reveal how the grasp patterns proposed in this thesis change with different configurations.

The findings presented in this thesis showed that only five grasps account for the majority of data in VR, suggesting that interaction designers aiming to achieve intuitive interaction modes at lower computational costs (such as for consumer-available applications) could focus on the main five grasps proposed in this thesis. The full VR taxonomy of grasp types can be considered for improved grasp models in systems where intuitive and natural interactions are mandatory, such as in professional training and simulations. Together with the VR taxonomy, the categorisation methods could be used to identify the most intuitive grasp types for virtual objects of particular characteristics, or design the virtual environment to have objects that fit within one category of objects and therefore simplify the grasps needed for intuitive interactions. The categorisation methods could also be furthered to understand how other object properties can be categorised and connected to grasp patterns for existing virtual environments, future VR configurations or other forms of interactions with virtual objects outside grasping.

9.3 Review of Aim and Objectives

The primary aim of this work was to evaluate grasping patterns in VR and develop the first VR Taxonomy of Grasp Types and was achieved through the following objectives:

1. **Review and determine current trends in 3D hand interaction and real grasping research.** This was achieved in Chapter 2 and 3 where existing methods for 3D hand interaction and methods for analysing real grasping patterns as well as existing real grasp taxonomies were reviewed.
2. **Define a methodology for collecting grasping patterns in VR suitable for determining grasping trends and taxonomies.** This was achieved in Chapter 4 by reviewing current trends in HCI taxonomy development and methods for analysing grasping metrics in real environments. These methods were adapted to consider VR constraints and a methodology for collecting and analysing grasping patterns in VR was defined.
3. **Explore and quantify the differences and similarities between grasping real objects and grasping virtual objects.** This was achieved in Chapter 5 where grasping patterns for real objects were compared against grasping

patterns for virtual objects, reporting on similarities and differences in grasp metrics.

4. **Measure the impact of object characteristics and tasks on grasping metrics in VR.** This was achieved in Chapters 6 and 8 where grasp metrics were analysed and reported for various object characteristics (shape, stability and complexity) as well as translate and docking tasks.

5. **Evaluate differences in grasping approach based on visual cues for avatar and thermal feedback representation.** This was achieved in Chapter 7 by exploring how grasping patterns change for different visual cues for thermal feedback (hot, cold and empty mug conditions) as well as different 3D hand representations.

6. **Synthesize grasp instances in the first VR Taxonomy of Grasp Types.** This was achieved in Chapter 6 where the first VR Taxonomy of Grasp Types was proposed and furthered in Chapter 8 where the taxonomy was re-iterated in a more complete version that covers more virtual object characteristics.

7. **Define and synthesize grasp patterns and potential applications of the taxonomy for virtual environment object grasping work.** This was achieved in Chapters 5, 6,7 and 8 where grasping patterns were discussed for each user study presented in this thesis, as well as Chapter 9 which provides an overview of potential applications of the taxonomy for HCI work.

9.4 Limitations and Future Work

9.4.1 Protocol

This work explored grasping interactions while participants were seated, however future work should consider how these patterns change for standing or more free walking natural interactions. Moreover, throughout the study, participants were not specifically instructed that grasping type, location and aperture were under study and were therefore free to interact with the virtual objects as they felt suitable. However, they could have experienced the *Good Subject Effect*. Notably, this is found when participants can respond to an experiment in ways that they believe confirm the hypothesis of the study (Nichols & Maner, 2008). This might have also influenced users to grasp virtual objects in a more realistic way, although this effect has been previously explored for virtual interaction and is supported by the reality-based interaction theory (Jacob et al., 2008). Although the methodology used in this thesis tried to mitigate this effect, it cannot be ruled out completely. Therefore, future work should consider conducting an elicitation study where users are allowed to interact freely in the virtual environment, without being asked to grasp objects in an intuitive way. The results could then be compared against the taxonomy presented in this thesis to further assess current results.

9.4.2 Referents

The taxonomies presented in this thesis were created using grasp information collected from 16 different objects as referents, chosen from the *Yale-Carnegie Mellon University-Berkeley Object and Model Set* (Calli et al., 2015). While these objects covered various aspects of the manipulation problem such as variety of shapes, sizes and textures, adding new object instances that can be categorised using the proposed methods, and linked to the grasps presented in this thesis will strengthen the body of work around VR human grasping and its applicability to various domains. Tasks were also used as referents in this work, however the only tasks explored were simple translate tasks and mixed docking tasks (translation and rotation). Exploring more complex tasks such as using specific tools (writing, using a pair of scissors) would provide a clearer overview on whether tasks in general influence grasping patterns in VR, as in real grasping literature (Napier, 1956; Feix et al., 2014a).

9.4.3 Visual cues

To understand how visual cues that substitute haptic feedback influence grasping patterns in VR, this work explored visual cues for thermal representations. However, this work only explored thermal visual cues on one virtual object, a mug and the applicability of these findings on other virtual objects is unknown. Future work should therefore look into how visual thermal cues influence grasping patterns for other virtual objects. Moreover, exploring visual cues for other types of haptic feedback such as texture and weight would provide more insights into how current training and simulation systems should improve their interaction paradigms to provide natural and intuitive grasping in VR.

9.4.4 Grasps

The work presented in this thesis analysed grasp metrics currently used for analysing grasping in real environments such as grasp type, location and dimension (Feix et al., 2014b, 2014a) and immersive environments such as grasp aperture (Al-Kalbani et al., 2016a). While these metrics provided insights for how users grasp virtual objects, future work might consider exploring other grasp metrics such as grasp displacement (Al-Kalbani et al., 2016a), which measures the position accuracy of the user's hands against the virtual object and might be used for informing collider design decisions together with the metrics presented in this thesis.

This work focused on analysing grasp metrics at the grasping point (load phase), however future work might consider analysing grasps during the transition phase for designing robust continual grasping interactions. Moreover, work presented in this thesis focused on one-hand interactions only, however, to provide a more complete overview of grasping patterns in VR, future work should consider looking at bimanual grasping, particularly if exploring interaction with larger objects (Piumsomboon et al., 2013).

9.4.5 Taxonomy

The taxonomy presented in this thesis was developed based on current grasping analysis methods, providing an overview of grasp types for virtual object characteristics. Future work should consider using this taxonomy as a basis and analysing grasping patterns in VR for other influencing factors such as other categorisation methods or different tasks. The application of these results for improved grasping experience against a benchmark grasp model could also be considered in future work, to determine the usability improvements for VR interaction. Finally, considering advances in XR technology, it would be interesting to investigate the transferability of this work into XR to understand the influencing factors between virtual and real object interaction in other immersive environments.

9.4.6 COVID-19

This work took place throughout the period of the COVID-19 pandemic. All data collected was prior to the pandemic or following COVID-19 safe regulations. While further evaluations could be taken to synthesize these results into a more usable system, unfortunately restrictions due to wearable devices for VR constrained these which should be explored in future work.

A | Ethics - Consent Form

A.1 General Information

Thank you for taking the time to complete this test. This study aims at analysing and understanding the act of grasping virtual objects. The study is a part of the PhD of a student, Andreea Dalia Blaga, located at the DMT Lab in Birmingham City University. Information regarding the position of the palm and the fingers, the interaction and the time to complete each task will be stored for further analysis. All data is anonymous.If you have any questions or require any further information regarding the test, please do not hesitate to contact me. Please fill in your details below. Once you have completed the next section, please move on to reading an overview of the test you are about to undertake. Please read and sign the consent form.

A.2 Details

Please fill in the blanks. An ID will be assigned to you. ID= First three letters of your surname + age + first three letters of your name. ID: .... Gender: .... Occupation: ....

Please answer the questions below:

- Have you been trained on this test before? (Yes/No)
- Are you familiar with computer games? (Novice, Intermediate, Advanced)
- Which of the below best describes your level of experience with virtual environments? (Novice, Intermediate, Advanced)

- Which of the below best describes your level of experience with gesture recognition systems? (Novice, Intermediate, Advanced)

A.3 Test Brief

You will be immersed into a virtual environment comprising a set of virtual objects. For each object you will be asked to grasp the object in the way you feel most intuitive, inform the test coordinator that you are ready to grasp, and move the object to the position of the target. The study will take [variation of minutes for each separate experiment] minutes.

B | Chapter 7 - Survey responses

Responses to the open-ended question (Do you think temperature representations influenced your interaction with the mug? Why?) presented in visual cues user experiment are:

- Because the one with steam and winter background made me change the way I grasp the mug in a sense of it created a cosy environment, so my hand wanted to be closer to the mug
- Yes- as it was a robotic hand I felt I could grab the mug with hot content from the top because I knew I won't get burnt
- Yes, natural reactions to avoid being burnt
- Yes, natural reactions to avoid being burnt
- yes. I was afraid not being burnt
- I was afraid of being burnt
- Yes, I was more careful when grabbing the hot mug and the the cold one when it had a warning and the subsequent ones with ice.
- Yes, I was more careful when I was grabbing the hot or cold mug.
- Yes. In real life I would have grasped the handle of the mug if it contained hot/cold fluids, and this is what I felt like doing here as well.
- Yes. I felt the need to grasp the handle of the mug when this contained hot/cold fluids, just like i would have done in real life.
- Yes. I used the handle for the warm mug and the main body of the mug for the iced water.

- Yes. The temperature representation influenced how to grasp the mug, as they do in the real world.
- No, I'm too used to the concept of VR being only visual to be afraid of hurting my fingers.
- No, I'm too used to VR applications being only visual to be afraid of hurting my hand.
- Yes I was slower with hot mugs
- Yes, I was more careful with hot mugs.
- More Cautious with the hot drinks than cold
- The Hotter ones seemed like they could burn me
- Yes. It affected the way I grabbed the mug. For example, with the "Hot" mug i was more careful in the way I grabbed it.
- Yes. It affected the way I grabbed the mug. Especially the Hot mug.
- Hot temperature influenced, in conjunction with the mug material so I did not burn my virtual hand yes, did not want to spill the hot coffee
- Because you feel like you don't want to grab a mug that is hot so you use the handle for your grip instead of the body of the mug
- You are less likely to grab something if you think it is going to burn you
- It did mildly due to being inside. I wasn't overly influenced by the surroundings
- There was influence once again, but not overly due to being an in inside environment.

- Yes, because when it was hot I was more careful in the way I was picking up the mug and when it was cold/empty I just picked it up quickly
- A little, as this time it was a robotics hand, therefore I felt a little less careful with how i was holding the mug with hot contents
- No, because I would always pick up a mug in the same way
- No, because I would always pick up a mug in the same way
- Yes ,because it looked like I was holding something hot or cold
- No
- No
- No
- No, I did not felt immersed to be influenced by that.
- I was cautious of the mug when steam was coming out of it as it indicates that it is hot
- Yes because it made me cautious about the temperature
- Yes, i tried to grab the mug in the right way
- Yes, because i tried to pick the mug in the right way and when it was full of coffee or water
- The different temperature representations affected the way I held the mug because I felt a need to be careful with the ’hotter’ mugs in a way that I didn’t feel with the empty or cold mugs
- The hot mugs made me not want to burn myself but the colder ones assured me it was alright to hold the mug in any way i wanted to

- I was trying to be more "careful" when grasping the handle if there was steam.
- When there was hot content in the mug I was keeping the mug closer to the table.
- Yes because if there was steam then it would indicate it was hot and wouldnt want to burn myself
- Yes because it helped decide how i would hold the cup - so i wouldnt get burned or cold hands
- If i see that the mug is hot i am more likely to pick it up from the handle
- Yes it help you decide where to grab the mug, if it is hot
- Yes
- Yes cause of the steam
- More caution with hot items, than cold or empty mugs
- More caution with mugs that had liquid
- Not really. Because I was aware that it is in virtual world
- No
- because i could see the steam
- the label
- based on the temperature the the location of the environment was changed and it made easier for me to assume the interaction
- Yes, based on the outside environment, I was influence by the interaction

- Yes, because it told me the warning sign that is it hot and the steam coming from it.
- Yes, felt like i would burn myself if i touched the Hot Mug
- No. I would pick them all up the same way.
- Yes as the hand looked more realistic
- Yes the temperature influenced my decision. i believe that it did that as it a normal reaction, to have when interacting with hot and cold objects in real life.
- The temperature did not affect my decision. As the robotic hand does not require too, much thinking about how to grasp the mug.
- yes, help me decide how to grasp it
- yes, on the way I grasp the mug
- because I was careful not to get burnt
- I did not want to get burnt
- yes, because the hot one might hurt me
- no, because i know it's not real
- yes. because it makes you pause and think rather than instinct
- Yes. Because it indicates the temperature
- Not too much.
- Not too much.
- No
- No

- Due to it feeling real, it influenced my interaction because i knew if it was hot or cold.
- Yeah due to me knowing the temperature.
- Yes, hot contents made me more wary of burning my hand
- Sort of. Due to having a robotic hand, the temperature of the contents of the mug didn't really matter to me as I couldn't burn my hand.
- Only hot temperature influenced my grasp because instinctively I didn't want to spill the content on me.
- When it was hot I was careful not to spill the coffee on me.
- I knew to be slightly more careful in the way in which i held the hot mugs
- Yes as I knew to hold the mugs with content by the handle rather than byteh cup itself
- It didnt as I was aware it wasnt reality.
- Yes as I got more scared and was more cautions to prevent injury.
- I'd normally grab a hot/cold object by the handle.
- No, the hand this time was robotic"
- yes becuase i didnt want to be burnt
- yes because i could see a human hand so i paid more attention to to the temperature
- No because i knew the temperature wouldnt affect me physically
- no because the temperature of the mug wouldn't affect me. also, holding the mug from the

- handle meant that i would't be affected by the temperatures
- No, as i knew the temperatures were not real so didn't hesitate to pick the mug up.
- No, as i knew the temperature wasn't real.
- The hotter/colder visualization of the the mug made me interact differently, by holding the handle so my virtual hand don't get burnt/cold.
- Yes as I was grabbing the mug differently based on temperature.
- No, even if I was drinking IRL from a mug, I would still use the handle.
- No, I grabbed by the handle for everything.
- Yes, when it seemed hot or cold I grabbed it by the handle
- No as it was a robot hand and wouldn't be affected by temprature

C | Chapter 8 - Survey responses

Responses to the open-ended question (Can you describe your strategy for grasping and placing the objects?) presented in visual cues user experiment are:

- I picked up objects from a specific angle so that I'd be able to rotate my hand to place them in the correct rotation. To do this I checked where the objects had to go before picking them up.
- I tried to first finish with the tool panel no matter where it was.
- I would tend to look at the target of some items first before grasping as depending on the position of the item on the table and the position of the item on the shelf I would decide to pick it up or not. It was dependent on if the interaction needed was more, so If i have to reposition the item on the table I would tend not to grasp these objects first.
- I usually selected objects that required to be further to me
- I was attracted to completing the area with the most missing objects first, and then looking at the closest objects to me that can go there.
- Grasping had a very high sensitivity and object was grasped even with the slightest twitch of my fingers rather than fully holding it
- The VR experience felt very much like the real world and completely influenced my decision regarding the way i grasp the object. I also took into account surrounding objects when placing an item into their position (i.e. the shelf when placing the item, the table when picking up an item)
- Generally I selected the object first and then looked for the target location. Where location or rotation were different that my intended grasp, I then

adjusted grasp position and location on the object.

- Start with the shelf and move from top left to bottom right. grasping objects from the top for fine positioning.
- focus on target object and then grasp firmly
- I would go for the easiest option first, then look at where it would need grasping.
- To begin with I just tried to grasp as I would normally in the real world. When placing I would get the object roughly into the correct area (sometimes correct on the first attempt) and then adjust to fit.
- I did not think about strategy I was just grabbing the objects to keep them in place
- Closer object first in the left handside of the table
- Deal with the nearest objects first
- I would first move the object to the target area roughly. and once in position I would then aim to rotate the object to match the target area and then once the objects rotation was correct, make fine adjustments to match the position more accurately
- I started with the shelf, grasped the objects in the order they had to be placed on the shelf, from left to right, from upper shelves to lower shelves. Continued in the same manner with the tool panel, and then the basket.
- first aim to what is in front of me, 2nd what is bigger, 3rd what object's rotation matched the target's one
- look at the object first then pick it after

- the objects which were closer to me were first preferred and then it went in random order
- Picked object in front of me at random, localised the target and then moved object to target.
- I chose the object that were closer to me. If I found targeted object in the table I would choose that.
- Nearer objects and target objects
- I would first look for the target and then grasp the object so that I can rotate and place it in the target position/rotation.

References

Achibet, M., Girard, A., Talvas, A., Marchal, M., & Lećuyer, A. (2015, March). Elastic-arm: Human-scale passive haptic feedback for augmenting interaction and perception in virtual environments. In *2015 ieee virtual reality (vr)* (p. 63-68). doi: 10.1109/VR.2015.7223325

Adamides, G., Christou, G., Katsanos, C., Xenos, M., & Hadzilacos, T. (2015). Usability guidelines for the design of robot teleoperation: A taxonomy. *IEEE Transactions on Human-Machine Systems*, *45*(2), 256-262. doi: 10.1109/THMS.2014.2371048

Adamo-Villani, N., & Wilbur, R. (2007). An immersive game for k-5 math and science education. In *2007 11th international conference information visualization (iv '07)* (p. 921-924). doi: 10.1109/IV.2007.23

Agah, A., & Tanie, K. (1999). Taxonomy of research on human interactions with intelligent systems. In *Ieee smc'99 conference proceedings. 1999 ieee international conference on systems, man, and cybernetics (cat. no.99ch37028)* (Vol. 6, p. 965-970 vol.6). doi: 10.1109/ICSMC.1999.816683

Alce, G., Wallergard, M., & Hermodsson, K. (2015, 06). Wozard: A wizard of oz method for wearable augmented reality interaction-a pilot study. *Advances in Human-Computer Interaction*, *2015*. doi: 10.1155/2015/271231

Aldhaban, F., Daim, T., & Harmon, R. (2015, 08). Exploring the adoption and use of the smartphone technology in emerging regions: A literature review and hypotheses development. In (p. 2355-2370). doi: 10.1109/PICMET.2015.7273210

Ali, A. X., Morris, M. R., & Wobbrock, J. O. (2021, March). Distributed interaction design: Designing human-centered interactions in a time of social distancing. *Interactions*, *28*(2), 82-87. Retrieved from https://doi.org/

10.1145/3447790 doi: 10.1145/3447790

Al-Kalbani, M., Frutos-Pascual, M., & Williams, I. (2017). Freehand grasping in mixed reality: Analysing variation during transition phase of interaction. In *Proceedings of the 19th acm international conference on multimodal interaction* (p. 110-114). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3136755.3136776 doi: 10.1145/3136755.3136776

Al-Kalbani, M., Frutos-Pascual, M., & Williams, I. (2019). Virtual object grasping in augmented reality: Drop shadows for improved interaction. In *2019 11th international conference on virtual worlds and games for serious applications (vs-games)* (p. 1-8). doi: 10.1109/VS-Games.2019.8864596

Al-Kalbani, M., Williams, I., & Frutos-Pascual, M. (2016a). Analysis of medium wrap freehand virtual object grasping in exocentric mixed reality. In *2016 ieee international symposium on mixed and augmented reality (ismar)* (p. 84-93). doi: 10.1109/ISMAR.2016.14

Al-Kalbani, M., Williams, I., & Frutos-Pascual, M. (2016b). Improving freehand placement for grasping virtual objects via dual view visual feedback in mixed reality. In *Proceedings of the 22nd acm conference on virtual reality software and technology* (p. 279-282). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2993369.2993401 doi: 10.1145/2993369.2993401

Almeida, L., Lopes, E., YalÃ§inkaya, B., Martins, R., Lopes, A., Menezes, P., & Pires, G. (2019). Towards natural interaction in immersive reality with a cyber-glove. In *2019 ieee international conference on systems, man and cybernetics (smc)* (p. 2653-2658). doi: 10.1109/SMC.2019.8914239

Ameur, S., Ben Khalifa, A., & Bouhlel, M. S. (2020). Hand-gesture-based touchless exploration of medical images with leap motion controller. In *2020 17th* *international multi-conference on systems, signals devices (ssd)* (p. 6-11). doi: 10.1109/SSD49366.2020.9364244

Angelini, L., Carrino, F., Carrino, S., Caon, M., Khaled, O. A., Baumgartner, J., ... Mugellini, E. (2014). Gesturing on the steering wheel: A user-elicited taxonomy. In *Proceedings of the 6th international conference on automotive user interfaces and interactive vehicular applications* (p. 1-8). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2667317.2667414 doi: 10.1145/2667317.2667414

Anjo, M. d. S., Pizzolato, E. B., & Feuerstack, S. (2012). A real-time system to recognize static gestures of brazilian sign language (libras) alphabet using kinect. In *Proceedings of the 11th brazilian symposium on human factors in computing systems* (p. 259-268). Porto Alegre, BRA: Brazilian Computer Society.

Anupama, H. S., Usha, B. A., Madhushankar, S., Vivek, V., & Kulkarni, Y. (2021). Automated sign language interpreter using data gloves. In *2021 international conference on artificial intelligence and smart systems (icais)* (p. 472-476). doi: 10.1109/ICAIS50930.2021.9395749

Araujo, B., Jota, R., Perumal, V., Yao, J. X., Singh, K., & Wigdor, D. (2016). Snake charmer: Physically enabling virtual objects. In *Proceedings of the tei'16: Tenth international conference on tangible, embedded, and embodied interaction* (pp. 218–226).

Arbib, M. A. (1985). Coordinated control programs for movements of the hand. *Experimental Brain Research*, *10*, 111-129.

Arora, R., Kazi, R. H., Kaufman, D. M., Li, W., & Singh, K. (2019). Magicalhands: Mid-air hand gestures for animating in vr. In *Proceedings of the 32nd annual acm symposium on user interface software and technology*

(p. 463-477). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3332165.3347942 doi: 10.1145/3332165.3347942

Augestein, T., Kortemeyer, D., & L, G. (2021). Enhancing mirror therapy via scaling and shared control: a novel open-source virtual reality platform for stroke rehabilitation. In (p. 525-53). New York, NY, USA: Virtual Reality (2021). Retrieved from https://doi.org/10.1007/s10055-021-00593-4 doi: 10.1007/s10055-021-00593-4

Avila, L., & Bailey, M. (2016). Augment your reality. *IEEE Computer Graphics and Applications*, *36*(1), 6-7. doi: 10.1109/MCG.2016.17

Azizo, A. S. b., Mohamed, F. b., Siang, C. V., & Isham, M. I. M. (2020). Virtual reality 360 utm campus tour with voice commands. In *2020 6th international conference on interactive digital media (icidm)* (p. 1-6). doi: 10.1109/ICIDM51048.2020.9339665

Bai, H., Gao, L., El-Sana, J., & Billinghurst, M. (2013). Markerless 3D gesture-based interaction for handheld Augmented Reality interfaces. *2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013*(October 2013), 0–5.

Bai, H., Nassani, A., Ens, B., & Billinghurst, M. (2017). Asymmetric bimanual interaction for mobile virtual reality. In *Proceedings of the 27th international conference on artificial reality and telexistence and 22nd eurographics symposium on virtual environments* (p. 83-86). Goslar, DEU: Eurographics Association.

Balani, M., & Tümler, J. (2021, 07). Usability and user experience of interactions on vr-pc, hololens 2, vr cardboard and ar smartphone in a biomedical application. In (p. 275-287). doi: 10.1007/978-3-030-77599-5_20

Barkokebas, R., Ritter, C., Sirbu, V., Li, X., & Al-Hussein, M. (2019). Application of virtual reality in task training in the construction manufacturing industry. In *Isarc. proceedings of the international symposium on automation and robotics in construction* (Vol. 36, pp. 796–803).

Beer, R., Rieth, C., Tran, R., & Cook, M. (2017, 03). Framework for multi-human multi-robot interaction: Impact of operational context and team configuration on interaction task demands. In (p. 287-293).

Bellucci, A., Zarraonandia, T., Díaz, P., & Aedo, I. (2021). Welicit: A wizard of oz tool for vr elicitation studies. In C. Ardito et al. (Eds.), *Human-computer-interaction – interact 2021* (pp. 82–91). Cham: Springer International Publishing.

Bendz, P. (1974). Systematization of the grip of the hand in relation to finger motor systems. *Scandinavian Journal of Rehabilitation Medicine*, 158-165.

Benko, H., Holz, C., Sinclair, M., & Ofek, E. (2016). Normaltouch and texturetouch: High-fidelity 3d haptic shape rendering on handheld virtual reality controllers. In *Proceedings of the 29th annual symposium on user interface software and technology* (pp. 717–728). ACM. Retrieved from http://doi.acm.org/10.1145/2984511.2984526 doi: 10.1145/2984511.2984526

Bennett, C. A., & Rey, P. (1972). What's so hot about red? *Human Factors*, *14*(2), 149-154. Retrieved from https://doi.org/10.1177/001872087201400204 (PMID: 5022474) doi: 10.1177/001872087201400204

Bennett, K. M. B., & Castiello, U. (1994, 01). Reach to Grasp: Changes With Age. *Journal of Gerontology*, *49*(1), P1-P7. Retrieved from https://doi.org/10.1093/geronj/49.1.P1 doi: 10.1093/geronj/49.1.P1

Bharath, V., & Patil, R. (2018). Solid modelling interaction with sensors for virtual reality welding. In *Matec web of conferences* (Vol. 144, p. 01008).

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological review*, *94 2*, 115-147.

Biederman, I., Cooper, E., Hummel, J., & Fiser, J. (1993, 01). Geon theory as an account of shape recognition in mind, brain and machine. doi: 10.5244/C.7.18

Billinghurst, M., Piumsomboon, T., & Bai, H. (2014, 05). Hands in space: Gesture interaction with augmented-reality interfaces. *Computer Graphics and Applications, IEEE*, *34*, 77-80. doi: 10.1109/MCG.2014.8

Biocca, F. (1992, 02). Virtual reality technology: A tutorial. *Journal of Communication*, *42*, 23 - 72. doi: 10.1111/j.1460-2466.1992.tb00811.x

Biocca, F., Kim, J., & Choi, Y. (2001). Visual touch in virtual environments: An exploratory study of presence, multimodal interfaces, and cross-modal sensory illusions. *Presence: Teleoperators and Virtual Environments*, *10*(3), 247-265. Retrieved from https://doi.org/10.1162/105474601300343595 doi: 10.1162/105474601300343595

Blaga, A. D., Frutos-Pascual, M., Creed, C., & Williams, I. (2020). Too hot to handle: An evaluation of the effect of thermal visual representation on user grasping interaction in virtual reality. In *Proceedings of the 2020 chi conference on human factors in computing systems* (p. 1-16). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3313831.3376554

Blaga, A. D., Frutos-Pascual, M., Creed, C., & Williams, I. (2021a). Freehand grasping: An analysis of grasping for docking tasks in virtual reality. In *2021 ieee virtual reality and 3d user interfaces (vr)* (p. 749-758). doi: 10.1109/VR50410.2021.00102

Blaga, A. D., Frutos-Pascual, M., Creed, C., & Williams, I. (2021b). A grasp on reality: Understanding grasping patterns for object interaction in real and virtual environments. In *2021 ieee international symposium on mixed and augmented reality adjunct (ismar-adjunct)* (p. 391-396). doi: 10.1109/ISMAR-Adjunct54149.2021.00090

Blaga, A. D., Frutos-Pascual, M., Creed, C., & Williams, I. (2021c). Virtual object categorisation methods: Towards a richer understanding of object grasping for virtual reality. In *Proceedings of the 27th acm symposium on virtual reality software and technology.* New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3489849.3489875 doi: 10.1145/3489849.3489875

Boritz, J., & Booth, K. (1998). A study of interactive 6 dof docking in a computerised virtual environment. *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180)*, 139-146.

Borst, C., & Indugula, A. (2005). Realistic virtual grasping. In *Ieee proceedings. vr 2005. virtual reality, 2005.* (p. 91-98). doi: 10.1109/VR.2005.1492758

Bostan, B. (2006). Interaction in virtual reality. *4th International Symposium of Interactive Medial Design (ISIMD).*

Bowman, D. A., & Hodges, L. F. (1999). Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments. *J. Vis. Lang. Comput.*, *10*, 37-53.

Bowman, D. A., McMahan, R. P., & Ragan, E. D. (2012, September). Questioning naturalism in 3d user interfaces. *Commun. ACM*, *55*(9), 78–88. Retrieved from http://doi.acm.org/10.1145/2330667.2330687 doi: 10.1145/2330667.2330687

Bozzacchi, C., Volcic, R., & Domini, F. (2014, 12). Effect of visual and haptic feedback on grasping movements. *Journal of Neurophysiology*, *112*, 3189-3196. doi: 10.1152/jn.00439.2014

Buchmann, V., Nilsen, T., & Billinghurst, M. (2005). Interaction with partially transparent hands and objects. In *Proceedings of the sixth australasian conference on user interface - volume 40* (pp. 17–20). Retrieved from http://dl.acm.org/citation.cfm?id=1082243.1082246

Bullock, I., Feix, T., & Dollar, A. (2015, 08). Human precision manipulation workspace: Effects of object size and number of fingers used. In (Vol. 2015, p. 5768-5772). doi: 10.1109/EMBC.2015.7319703

Bullock, I., Zheng, J., Rosa, S., Guertler, C., & Dollar, A. (2013, 07). Grasp frequency and usage in daily household and machine shop tasks. *Haptics, IEEE Transactions on*, *6*, 296-308. doi: 10.1109/TOH.2013.6

Caggianese, G., Gallo, L., & Neroni, P. (2019, 01). The vive controllers vs. leap motion for interactions in virtual environments: A comparative evaluation. In (p. 24-33). doi: 10.1007/978-3-319-92231-7_3

Cai, M., Kitani, K. M., & Sato, Y. (2016). Understanding hand-object manipulation with grasp types and object attributes. In *Robotics: Science and systems.*

Calli, B., Walsman, A., Singh, A., Srinivasa, S., Abbeel, P., & Dollar, A. M. (2015, Sep.). Benchmarking in manipulation research: Using the yale-cmu-berkeley object and model set. *IEEE Robotics Automation Magazine*, *22*(3), 36-52. doi: 10.1109/MRA.2015.2448951

Cameron, C. R., DiValentin, L. W., Manaktala, R., McElhaney, A. C., Nostrand, C. H., Quinlan, O. J., ... Gerling, G. J. (2011). Hand tracking and visualization in a virtual reality simulation. In *2011 ieee systems and information engineering design symposium* (p. 127-132). doi: 10.1109/SIEDS.2011.5876867

Camp, F., Schick, A., & Stiefelhagen, R. (2013). How to click in mid-air. In *Proceedings of the first international conference on distributed, ambient, and pervasive interactions - volume 8028* (p. 78-86). Berlin, Heidelberg:

Springer-Verlag. Retrieved from https://doi.org/10.1007/978-3-642-39351-8_9 doi: 10.1007/978-3-642-39351-8_9

Cao, Q., Gao, C., Wang, Y., & Li, T. (2016). The motion capture data glove device for virtual surgery. In *2016 ieee international nanoelectronics conference (inec)* (p. 1-2). doi: 10.1109/INEC.2016.7589461

Capece, N., Erra, U., & Grippa, J. (2018). Graphvr: A virtual reality tool for the exploration of graphs with htc vive system. In *2018 22nd international conference information visualisation (iv)* (p. 448-453). doi: 10.1109/iV.2018.00084

Carfi, A., & Mastrogiovanni, F. (2021, 12). Gesture-based human-machine interaction: Taxonomy, problem definition, and analysis. *IEEE Transactions on Cybernetics*, *PP*, 1-17. doi: 10.1109/TCYB.2021.3129119

Carmeli, E., Patish, H., & Coleman, R. (2003, 03). The aging hand. *The journals of gerontology. Series A, Biological sciences and medical sciences*, *58*, 146-52. doi: 10.1093/gerona/58.2.M146

Cauchard, J. R., Tamkin, A., Wang, C. Y., Vink, L., Park, M., Fang, T., & Landay, J. A. (2019). Drone.io: A gestural and visual interface for human-drone interaction. In *2019 14th acm/ieee international conference on human-robot interaction (hri)* (p. 153-162). doi: 10.1109/HRI.2019.8673011

Cesari, P., & Newell, K. (1999, 09). The scaling of human grip configurations. *Journal of experimental psychology. Human perception and performance*, *25*, 927-35. doi: 10.1037//0096-1523.25.4.927

Cesari, P., & Newell, K. (2000, 11). Body-scaled transitions in human grip configurations. *Journal of experimental psychology. Human perception and performance*, *26*, 1657-68. doi: 10.1037/0096-1523.26.5.1657

Cesari, P., & Newell, K. (2002, 11). Scaling the components of prehension. *Motor control*, *6*, 347-65. doi: 10.1123/mcj.6.4.347

Chen, D., Liu, H., & Ren, Z. (2018). Application of wearable device htc vive in upper limb rehabilitation training. In *2018 2nd ieee advanced information management,communicates,electronic and automation control conference (imcec)* (p. 1460-1464). doi: 10.1109/IMCEC.2018.8469540

Chen, D. L., Balakrishnan, R., & Grossman, T. (2020). Disambiguation techniques for freehand object manipulations in virtual reality. In *2020 ieee conference on virtual reality and 3d user interfaces (vr)* (p. 285-292). doi: 10.1109/VR46266.2020.00048

Chen, T., Xu, L., Xu, X., & Zhu, K. (2021). Gestonhmd: Enabling gesture-based interaction on low-cost vr head-mounted display. *IEEE Transactions on Visualization and Computer Graphics*, *27*(5), 2597-2607. doi: 10.1109/TVCG.2021.3067689

Chen, Z., Peng, W., Peiris, R., & Minamizawa, K. (2017, 07). Thermoreality: thermally enriched head mounted displays for virtual reality. In (p. 1-2). doi: 10.1145/3102163.3102222

Chheang, V., Apilla, V., Saalfeld, P., Boedecker, C., Huber, T., Huettl, F., … Hansen, C. (2021). Collaborative vr for liver surgery planning using wearable data gloves: An interactive demonstration. In *2021 ieee conference on virtual reality and 3d user interfaces abstracts and workshops (vrw)* (p. 768-768). doi: 10.1109/VRW52623.2021.00268

Choi, I., Ofek, E., Benko, H., Sinclair, M., & Holz, C. (2018). Claw: A multi-functional handheld haptic controller for grasping, touching, and triggering in virtual reality. In *Proceedings of the 2018 chi conference on human factors in computing systems* (p. 1-13). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3173574.3174228

Coburn, J., Freeman, I., & Salmon, J. L. (2017). A review of the capabilities of current low-cost virtual reality technology and its potential to enhance the design process. *J. Comput. Inf. Sci. Eng.*, *17*(3).

Cohen, M. W., Voldman, I., Regazzoni, D., & Vitali, A. (2018). Hand rehabilitation via gesture recognition using leap motion controller. In *2018 11th international conference on human system interaction (hsi)* (p. 404-410). doi: 10.1109/HSI.2018.8431349

Cook, D. M., Dissanayake, D., & Kaur, K. (2019). Virtual reality and older hands: Dexterity and accessibility in hand-held vr control. In *Proceedings of the 5th international acm in-cooperation hci and ux conference* (p. 147-151). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3328243.3328262 doi: 10.1145/3328243.3328262

Cooke, L., & Cuddihy, E. (2005). Using eye tracking to address limitations in think-aloud protocol. In *Ipcc 2005. proceedings. international professional communication conference, 2005.* (p. 653-658). doi: 10.1109/IPCC.2005.1494236

Cooper, N., Milella, F., Pinto, C., Cant, I., White, M., & Meyer, G. (2018). The effects of substitute multisensory feedback on task performance and the sense of presence in a virtual reality environment. *PloS one*, *13*(2).

Covarrubias, M., & Bordegoni, M. (2015). Immersive vr for natural interaction with a haptic interface for shape rendering. In *2015 ieee 1st international forum on research and technologies for society and industry leveraging a better tomorrow (rtsi)* (p. 82-89). doi: 10.1109/RTSI.2015.7325075

Cridlin, L. (2007, 01). The importance of hands-on learning. In (p. 151-156). doi: 10.2351/1.5056625

Cutkosky, M., & Kao, I. (1989). Computing and controlling compliance of a robotic hand. *IEEE Transactions on Robotics and Automation*, *5*(2), 151-

165. doi: 10.1109/70.88036

Cutkosky, M., & Wright, P. (1986). Modeling manufacturing grips and correlations with the design of robotic hands. In *Robotics and automation. proceedings. 1986 ieee international conference on* (Vol. 3, pp. 1533–1539).

Cutkosky, M. R. (1989, Jun). On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Transactions on Robotics and Automation*, *5*(3), 269-279. doi: 10.1109/70.34763

Cutkosky, M. R., & Howe, R. D. (1990). Human Grasp Choice and Robotic Grasp Analysis. *Dextrous Robot Hands*, *1*, 5–31.

Cutkosky, M. R., & Wright, P. K. (1986). Friction, stability and the design of robotic fingers. *The International Journal of Robotics Research*, *5*(4), 20-37. Retrieved from https://doi.org/10.1177/027836498600500402 doi: 10.1177/027836498600500402

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of oz studies: Why and how. In *Proceedings of the 1st international conference on intelligent user interfaces* (pp. 193–200). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/169891.169968 doi: 10.1145/169891.169968

Datcu, D., & Lukosch, S. (2013). Free-hands interaction in augmented reality. In *Proceedings of the 1st symposium on spatial user interaction* (p. 33-40). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2491367.2491370 doi: 10.1145/2491367.2491370

Derbyshire, N., Ellis, R., & Tucker, M. (2006). The potentiation of two components of the reach-to-grasp action during object categorisation in visual memory. *Acta Psychologica*, *122*(1), 74-98. Retrieved from https://www.sciencedirect.com/science/article/

pii/S0001691805001174 doi: https://doi.org/10.1016/j.actpsy.2005.10.004

Dhanasree, K., Nisha, K., & Jayakrishnan, R. (2018). Hospital emergency room training using virtual reality and leap motion sensor. In *2018 second international conference on intelligent computing and control systems (iciccs)* (p. 924-928). doi: 10.1109/ICCONS.2018.8662900

Dinh, H. Q., Walker, N., Hodges, L. F., Song, C., & Kobayashi, A. (1999, March). Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings ieee virtual reality (cat. no. 99cb36316)* (p. 222-228). doi: 10.1109/VR.1999.756955

Dipietro, L., Sabatini, A. M., & Dario, P. (2008). A survey of glove-based systems and their applications. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *38*(4), 461-482. doi: 10.1109/TSMCC.2008.923862

Edsinger, A., & Kemp, C. C. (2007). Human-robot interaction for cooperative manipulation: Handing objects to one another. *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*, 1167-1172.

Elangovan, N., Chang, C.-M., Gao, G., & Liarokapis, M. (2022, 04). An accessible, open-source dexterity test: Evaluating the grasping and dexterous manipulation capabilities of humans and robots. *Frontiers in Robotics and AI*, *9*, 808154. doi: 10.3389/frobt.2022.808154

El-Shimy, D., Marentakis, G., & Cooperstock, J. (2009, 01). Tech-note: Multimodal feedback in 3d target acquisition. In (p. 95-98). doi: 10.1109/3DUI.2009.4811212

Englmeier, D., Dörner, J., Butz, A., & Höllerer, T. (2020). A tangible spherical proxy for object manipulation in augmented reality. In *2020 ieee conference* *on virtual reality and 3d user interfaces (vr)* (p. 221-229).

Ericson, C. (2004). *Real-time collision detection*. USA: CRC Press, Inc.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis: Verbal reports as data..

Farid, R., & Sammut, C. (2013, 04). Plane-based object categorisation using relational learning. *Machine Learning*, 1-21. doi: 10.1007/s10994-013-5352-9

Feix, T., Bullock, I., & Dollar, A. (2014a, 10). Analysis of human grasping behavior: Correlating tasks, objects and grasps. *Haptics, IEEE Transactions on*, *7*, 430-441. doi: 10.1109/TOH.2014.2326867

Feix, T., Bullock, I., & Dollar, A. (2014b, 07). Analysis of human grasping behavior: Object characteristics and grasp type. *Haptics, IEEE Transactions on*, *7*, 311-323. doi: 10.1109/TOH.2014.2326871

Feix, T., Bullock, I. M., & Dollar, A. M. (2014c, July). Analysis of human grasping behavior: Object characteristics and grasp type. *IEEE Transactions on Haptics*, *7*(3), 311-323. doi: 10.1109/TOH.2014.2326871

Feix, T., Pawlik, R., & Schmiedmayer, H.-B. (2009, 01). A comprehensive grasp taxonomy. *Robotics, Science and Systems Conference: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation*.

Feix, T., Romero, J., Schmiedmayer, H. B., Dollar, A. M., & Kragic, D. (2016, Feb). The grasp taxonomy of human grasp types. *IEEE Transactions on Human-Machine Systems*, *46*(1), 66-77.

Flanagan, J. R., & Johansson, R. S. (2001). Hand movements. Encyclopedia of the Human Brain.

Fong, K., Tang, Y., Sie, K., Yu, A., Lo, C., & Ma, Y. (2021). Task-specific virtual reality training on hemiparetic upper extremity in patients with stroke. *Virtual Reality*. (Publisher Copyright: © 2021, The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature.) doi: 10.1007/s10055-021-00583-6

Franco, J., & Cabral, D. (2019). Augmented object selection through smart glasses. In *Proceedings of the 18th international conference on mobile and ubiquitous multimedia* (p. 1-5). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3365610.3368416 doi: 10.1145/3365610.3368416

Friedman, M. (1940). A comparison of alternative tests of significance for the problem of m rankings. *The Annals of Mathematical Statistics*, *11*(1), 86–92.

Fröhlich, J., & Wachsmuth, I. (2013, 07). The visual, the auditory and the haptic- a user study on combining modalities in virtual worlds. In (Vol. 8021, p. 159-168). doi: 10.1007/978-3-642-39405-8_19

Frutos-Pascual, M., Harrison, J. M., Creed, C., & Williams, I. (2019). Evaluation of ultrasound haptics as a supplementary feedback cue for grasping in virtual environments. In *2019 international conference on multimodal interaction* (pp. 310–318).

Fu, Q., Fu, J., Guo, J., Guo, S., & Li, X. (2020). Gesture recognition based on bp neural network and data glove. In *2020 ieee international conference on mechatronics and automation (icma)* (p. 1918-1922). doi: 10.1109/ICMA49215.2020.9233629

Furmanek, M. P., Mangalam, M., Lockwood, K., Smith, A., Yarossi, M., & Tunik, E. (2021). Effects of sensory feedback and collider size on reach-to-grasp coordination in haptic-free virtual reality. *Frontiers in Virtual Reality*, *2*. Retrieved from https://www.frontiersin.org/articles/10.3389/frvir.2021.648529 doi: 10.3389/frvir.2021.648529

Furmanek, M. P., Schettino, L. F., Yarossi, M., Kirkman, S., Adamovich, S. V.,

& Tunik, E. (2019). Coordination of reach-to-grasp in physical and haptic-free virtual environments. *Journal of NeuroEngineering and Rehabilitation*, *16*.

Gabbard, J. L. (1997). A taxonomy of usability characteristics in virtual environments. ETDs: Virginia Tech Electronic Theses and Dissertations.

Gadekallu, T., Alazab, M., Kaluri, R., Reddy, P., Bhattacharya, S., Lakshman, K., & M, P. (2021, 03). Hand gesture classification using a novel cnn-crow search algorithm. *Complex and Intelligent Systems*, *7*. doi: 10.1007/s40747-021-00324-x

García Álvarez, A., Roby-Brami, A., Robertson, J., & Roche, N. (2017, 11). Functional classification of grasp strategies used by hemiplegic patients. *PLOS ONE*, *12*(11), 1-21. Retrieved from https://doi.org/10.1371/journal.pone.0187608 doi: 10.1371/journal.pone.0187608

Ghomi, E., Huot, S., Bau, O., Beaudouin-Lafon, M., & Mackay, W. E. (2013). Arpège: Learning multitouch chord gestures vocabularies. In *Proceedings of the 2013 acm international conference on interactive tabletops and surfaces* (p. 209-218). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2512349.2512795 doi: 10.1145/2512349.2512795

Goldfeder, C., Ciocarlie, M., Dang, H., & Allen, P. K. (2009). The columbia grasp database. In *2009 ieee international conference on robotics and automation* (p. 1710-1716). doi: 10.1109/ROBOT.2009.5152709

Gonzalez, D. C., & Garnique, L. V. (2018). Development of a simulator with htc vive using gamification to improve the learning experience in medical students. In *2018 congreso internacional de innovacion y tendencias en ingenieria (coniiti)* (p. 1-6). doi: 10.1109/CONIITI.2018.8587058

Good, M. D., Whiteside, J. A., Wixon, D. R., & Jones, S. J. (1984, October). Building a user-derived interface. *ACM*, *27*(10), 1032-1043. Retrieved from https://doi.org/10.1145/358274.358284 doi: 10.1145/358274.358284

Griffiths, H. E. (1943). Treatment of the injured workman. *The Lancet*, *241*, 729-733.

Guan, H., & Zhang, J. (2016). Multi-sensory based novel household object categorization system by using interactive behaviours. In *2016 ieee international conference on robotics and biomimetics (robio)* (p. 1685-1690). doi: 10.1109/ROBIO.2016.7866570

Guzsvinecz, T., Szucs, V., & Sik-Lanyi, C. (2019). Suitability of the kinect sensor and leap motion controller-a literature review. *Sensors*, *19*(5). Retrieved from https://www.mdpi.com/1424-8220/19/5/1072 doi: 10.3390/s19051072

Hameed, S., Khan, M. A., Kumar, B., Arain, Z., & Hassan, M. (2017, 12). Gesture controlled robotic arm using leap motion. *Indian Journal of Science and Technology*, *10*, 1-7. doi: 10.17485/ijst/2017/v10i45/120630

Han, Y. (2010). A low-cost visual motion data glove as an input device to interpret human hand gestures. *IEEE Transactions on Consumer Electronics*, *56*(2), 501-509. doi: 10.1109/TCE.2010.5505962

Harley, D., Verni, A., Willis, M., Ng, A., Bozzo, L., & Mazalek, A. (2018). Sensory vr: Smelling, touching, and eating virtual reality. In *Proceedings of the twelfth international conference on tangible, embedded, and embodied interaction* (pp. 386–397). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/3173225.3173241 doi: 10.1145/3173225.3173241

Hart, S. G., & Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In P. A. Hancock &

N. Meshkati (Eds.), *Human mental workload* (Vol. 52, p. 139 - 183). North-Holland. doi: https://doi.org/10.1016/S0166-4115(08)62386-9

Hartney, J. H., Rosenthal, S. N., Kirkpatrick, A. M., Skinner, J. M., Hughes, J., & Orlosky, J. (2019). Revisiting virtual reality for practical use in therapy: Patient satisfaction in outpatient rehabilitation. In *2019 ieee conference on virtual reality and 3d user interfaces (vr)* (p. 960-961). doi: 10.1109/VR.2019.8797857

Heim, M. (2000). *Virtual realism.* Oxford University Press USA.

Hertel, J., Karaosmanoglu, S., Schmidt, S., Braeker, J., Semmann, M., & Steinicke, F. (2021). A taxonomy of interaction techniques for immersive augmented reality based on an iterative literature review. In *2021 ieee international symposium on mixed and augmented reality (ismar)* (p. 431-440). doi: 10.1109/ISMAR52148.2021.00060

Hock, P., Benedikter, S., Gugenheimer, J., & Rukzio, E. (2017). Carvr: Enabling in-car virtual reality entertainment. In *Proceedings of the 2017 chi conference on human factors in computing systems* (p. 4034-4044). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3025453.3025665 doi: 10.1145/3025453.3025665

Hoff, L., Hornecker, E., & Bertel, S. (2016). Modifying gesture elicitation: Do kinaesthetic priming and increased production reduce legacy bias? In *Proceedings of the tei '16: Tenth international conference on tangible, embedded, and embodied interaction* (p. 86-91). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2839462.2839472 doi: 10.1145/2839462.2839472

Höll, M., Oberweger, M., Arth, C., & Lepetit, V. (2018). Efficient physics-based implementation for realistic hand-object interaction in virtual reality. In

2018 ieee conference on virtual reality and 3d user interfaces (vr) (p. 175-182). doi: 10.1109/VR.2018.8448284

Holz, D., Ullrich, S., Wolter, M., & Kuhlen, T. (2008). Multi-contact grasp interaction for virtual environments. *JVRB - Journal of Virtual Reality and Broadcasting*, *5(2008)*(7).

Howard, W., & Kumar, V. (1996). On the stability of grasped objects. *IEEE Transactions on Robotics and Automation*, *12*(6), 904-917. doi: 10.1109/70.544773

Huang, D.-A., Ma, M., Ma, W.-C., & Kitani, K. M. (2015). How do we use our hands? discovering a diverse set of common grasps. In *2015 ieee conference on computer vision and pattern recognition (cvpr)* (p. 666-675). doi: 10.1109/CVPR.2015.7298666

Hudson, S., Matson-Barkat, S., Pallamin, N., & Jegou, G. (2018). With or without you? interaction and immersion in a virtual reality experience. *Journal of Business Research*. Retrieved from http://www.sciencedirect.com/science/article/pii/S0148296318305514 doi: https://doi.org/10.1016/j.jbusres.2018.10.062

Hudson, S., Matson-Barkat, S., Pallamin, N., & Jegou, G. (2019). With or without you? interaction and immersion in a virtual reality experience. *Journal of Business Research*, *100*, 459-468. Retrieved from https://www.sciencedirect.com/science/article/pii/S0148296318305514 doi: https://doi.org/10.1016/j.jbusres.2018.10.062

Hummels, C., & Stappers, P. (1998). Meaningful gestures for human computer interaction: beyond hand postures. In *Proceedings third ieee international conference on automatic face and gesture recognition* (p. 591-596). doi: 10.1109/AFGR.1998.671012

Iberall, T. (1987). The nature of human prehension: Three dextrous hands in one. In *Proceedings. 1987 ieee international conference on robotics and automation* (Vol. 4, p. 396-401). doi: 10.1109/ROBOT.1987.1087813

Iberall, T., Bingham, G. P., & Arbib, M. A. (1986). Opposition space as a structuring concept for the analysis of skilled hand movements..

II, J. E. S., Kuparinen, L., Rapson, S., & Sandor, C. (2017). Visually perceived distance judgments: Tablet-based augmented reality versus the real world. *International Journal of Human-Computer Interaction*, *33*(7), 576-591. Retrieved from https://doi.org/10.1080/10447318.2016.1265783 doi: 10.1080/10447318.2016.1265783

Islam, M. S., & Lim, S. (2022, 05). Vibrotactile feedback in virtual motor learning: A systematic review. *Applied Ergonomics*, *101*. doi: 10.1016/j.apergo.2022.103694

Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008). Reality-based interaction: A framework for post-wimp interfaces. In *Proceedings of the sigchi conference on human factors in computing systems* (p. 201-210). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1357054.1357089 doi: 10.1145/1357054.1357089

Jacobs, J., & Froehlich, B. (2011). A soft hand model for physically-based manipulation of virtual objects. In *2011 ieee virtual reality conference* (p. 11-18). doi: 10.1109/VR.2011.5759430

Jain, D., Junuzovic, S., Ofek, E., Sinclair, M., Porter, J., Yoon, C., . . . Ringel Morris, M. (2021). A taxonomy of sounds in virtual reality. In *Designing interactive systems conference 2021* (p. 160-170). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3461778.3462106 doi: 10.1145/3461778.3462106

Jayaram, S., Vance, J., Gadh, R., Jayaram, U., & Srinivasan, H. (2001, mar). Assessment of vr technology and its applications to engineering problems. *J. Comput. Info. Sci. Eng.*, *1*(1), 72-83. Retrieved from https://doi.org/10.1115/1.1353846 doi: 10.1115/1.1353846

J. Gianaros, P., Muth, E., Mordkoff, J., Levine, M., & M. Stern, R. (2001, 02). A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, space, and environmental medicine*, *72*, 115-9.

Jonas, M., Said, S., Yu, D., Aiello, C., Furlo, N., & Zytko, D. (2019). Towards a taxonomy of social vr application design. In *Extended abstracts of the annual symposium on computer-human interaction in play companion extended abstracts* (p. 437-444). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3341215.3356271 doi: 10.1145/3341215.3356271

Kamakura, N., Matsuo, M., Ishii, H., Mitsuboshi, F., & Miura, Y. (1980, 08). Patterns of static prehension in normal hands. *The American journal of occupational therapy : official publication of the American Occupational Therapy Association*, *34*, 437-45.

Kamal, M. M., & Sultana, K. Z. (2012). A comprehensive tool for text categorization and text summarization in bioinformatics. In *2012 15th international conference on computer and information technology (iccit)* (p. 592-597). doi: 10.1109/ICCITechn.2012.6509764

Kapandji, I., Honoré, L., & Poilleux, F. (1982). *The physiology of the joints: annotated diagrams of the mechanics of the human joints. upper limb*. Churchill Livingstone. Retrieved from https://books.google.at/books?id=DerdnAEACAAJ

Karam, M., & Schraefel, M. C. (2005). A study on the use of semaphoric gestures to support secondary task interactions. In *Chi '05 extended ab-*

stracts on human factors in computing systems (p. 1961-1964). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1056808.1057067 doi: 10.1145/1056808.1057067

Kawabe, T., Ujitoko, Y., & Yokosaka, T. (2022, 05). The relationship between illusory heaviness sensation and the motion speed of visual feedback in gesture-based touchless inputs. *Frontiers in Psychology*, *13*. doi: 10.3389/fpsyg.2022.811881

Keller, C. L. T., A. D, & Zahm, V. (1947). *Studies to determine the functional requirements for hand and arm prosthesis*. Department of Engineering, University of California at Los Angeles.

Kelley, J. (1985). *Cal - a natural language program developed with the oz paradigm: Implications for super systems*. IBM Thomas J. Watson Research Center. Retrieved from https://books.google.at/books?id=gNgaHQAACAAJ

Kerr, J., & Roth, B. (1986). Analysis of multifingered hands. *The International Journal of Robotics Research*, *4*(4), 3-17. Retrieved from https://doi.org/10.1177/027836498600400401 doi: 10.1177/027836498600400401

Kerzel, M., Ali, M., Ng, H. G., & Wermter, S. (2017, 05). Haptic material classification with a multi-channel neural network. In (p. 439-446). doi: 10.1109/IJCNN.2017.7965887

Khenak, N., Vézien, J., & Bourdot, P. (2020). Spatial presence, performance, and behavior between real, remote, and virtual immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, *26*(12), 3467-3478. doi: 10.1109/TVCG.2020.3023574

Kilteni, K., Groten, R., & Slater, M. (2012, dec). The sense of embodiment in virtual reality. *Presence: Teleoper. Virtual Environ.*, *21*(4), 373-387. Retrieved from https://doi.org/10.1162/PRES_a_00124 doi: 10.1162/PRES_a_00124

Kim, K., Park, M., & Lim, Y.-k. (2021). Guiding preferred driving style using voice in autonomous vehicles: An on-road wizard-of-oz study. In *Designing interactive systems conference 2021* (p. 352-364). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3461778.3462056 doi: 10.1145/3461778.3462056

Kim, M., & Lee, J. Y. (2016, dec). Touch and hand gesture-based interactions for directly manipulating 3d virtual objects in mobile augmented reality. *Multimedia Tools Appl.*, *75*(23), 16529-16550. Retrieved from https://doi.org/10.1007/s11042-016-3355-9 doi: 10.1007/s11042-016-3355-9

Kim, Y. R., Choi, H., Chang, M., & Kim, G. J. (2020). Applying touchscreen based navigation techniques to mobile virtual reality with open clip-on lenses. *Electronics*, *9*(9). Retrieved from https://www.mdpi.com/2079-9292/9/9/1448 doi: 10.3390/electronics9091448

Kipp, M., Neff, M., & Albrecht, I. (2007, 12). An annotation scheme for conversational gestures: How to economically capture timing and form. *Language Resources and Evaluation*, *41*, 325-339. doi: 10.1007/s10579-007-9053-5

Kleinman, E., Preetham, N., Teng, Z., Bryant, A., & Seif El-Nasr, M. (2021, oct). "what happened here!?" a taxonomy for user interaction with spatio-temporal game data visualization. *Proc. ACM Hum.-Comput. Interact.*, *5*(CHI PLAY). Retrieved from https://doi.org/10.1145/3474687 doi: 10.1145/3474687

Klompmaker, F., Paelke, V., & Fischer, H. (2013). A taxonomy-based approach towards nui interaction design. In *Proceedings of the first international conference on distributed, ambient, and pervasive interactions - volume 8028* (p. 32-41). Berlin, Heidelberg: Springer-Verlag. Retrieved from https://doi.org/10.1007/978-3-642-39351-8_4 doi: 10.1007/978-3-642-39351-8_4

Knierim, P., Schwind, V., Feit, A. M., Nieuwenhuizen, F., & Henze, N. (2018). Physical keyboards in virtual reality: Analysis of typing performance and effects of avatar hands. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 345:1–345:9). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/3173574.3173919 doi: 10.1145/3173574.3173919

Kreimeier, J., & Gotzelmann, T. (2020, 11). Two decades of touchable and walkable virtual reality for blind and visually impaired people: A high-level taxonomy. *Multimodal Technologies and Interaction*, *4*, 21. doi: 10.3390/mti4040079

Kristan, W., Skalak, R., Wilson, R., Skierczynski, B., Murray, J., Eisenhart, F., ... Beer, R. (2000, 01). Biomechanics and neural control of posture and movement. In (p. 206-220). doi: 10.1007/978-1-4612-2104-3_14

Kwasnik, B. (1999, 06). The role of classification in knowledge representation and discovery. *Library Trends*, *48*.

Lages, W. S., Nabiyouni, M., & Arantes, L. (2016). Krinkle cube: A collaborative vr game using natural interaction. In *Proceedings of the 2016 annual symposium on computer-human interaction in play companion extended abstracts* (p. 189-196). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2968120.2987746 doi: 10.1145/2968120.2987746

Landsmeer, J. (1962). Power grip and precision handling. *Annals of the rheumatic* *diseases*, *21*(2), 164.

Lawson, G., Roper, T., Shaw, E., Hsieh, M. K., & Cobb, S. V. (2020). Multimodal virtual environments: an opportunity to improve fire safety training? *Policy and Practice in Health and Safety*, *18*(2), 155-168. Retrieved from https://doi.org/10.1080/14773996.2020.1796085 doi: 10.1080/14773996.2020.1796085

Lazaro, M. J., Kim, S., Lee, J., Chun, J., Kim, G., Yang, E., ... Yun, M. (2021, 07). A review of multimodal interaction in intelligent systems. In (p. 206-219). doi: 10.1007/978-3-030-78462-1_15

Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, *19*(3), 342 - 368. Retrieved from http://www.sciencedirect.com/science/article/pii/0010028587900089 doi: https://doi.org/10.1016/0010-0285(87)90008-9

Lee, J., Sinclair, M., Gonzalez-Franco, M., Ofek, E., & Holz, C. (2019). Torc: A virtual reality controller for in-hand high-dexterity finger interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems* (p. 1-13). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3290605.3300301 doi: 10.1145/3290605.3300301

Lee, M., & Billinghurst, M. (2008). A wizard of oz study for an ar multimodal interface. In *Proceedings of the 10th international conference on multimodal interfaces* (p. 249-256). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1452392.1452444 doi: 10.1145/1452392.1452444

Leibe, B., & Schiele, B. (2003). Analyzing appearance and contour based methods for object categorization. In *2003 ieee computer society conference on* *computer vision and pattern recognition, 2003. proceedings.* (Vol. 2, p. II-409). doi: 10.1109/CVPR.2003.1211497

Levin, M., Magdalon, E., Michaelsen, S., & Quevedo, A. (2008, 09). Comparison of reaching and grasping kinematics in patients with hemiparesis and in healthy controls in virtual and physical environments. In (p. 60 -60). doi: 10.1109/ICVR.2008.4625123

Levin, M., Magdalon, E., Michaelsen, S., & Quevedo, A. (2015, 01). Quality of grasping and the role of haptics in a 3d immersive virtual reality environment in individuals with stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *23*. doi: 10.1109/TNSRE.2014.2387412

Li, B., Lu, Y., Li, C.-F., Godil, A., Schreck, T., Aono, M., . . . Tatsuma, A. (2014, 04). Large scale comprehensive 3d shape retrieval.. doi: 10.2312/3dor.20141059

Li, Y., Huang, J., Tian, F., Wang, H.-A., & Dai, G.-Z. (2019). Gesture interaction in virtual reality. *Virtual Reality and Intelligent Hardware*, *1*(1), 84-112. Retrieved from https://www.sciencedirect.com/science/article/pii/S2096579619300075 doi: https://doi.org/10.3724/SP.J.2096-5796.2018.0006

Lin, L., & Jörg, S. (2016). Need a hand?: How appearance affects the virtual hand illusion. In *Proceedings of the acm symposium on applied perception* (pp. 69–76). Retrieved from http://doi.acm.org/10.1145/2931002.2931006 doi: 10.1145/2931002.2931006

Liu, Y., Jiang, L., Liu, H., & Ming, D. (2022, 02). A straightforward and miniature implementation method of postural synergies for replicating human grasp characteristics accurately and intuitively. *Bioinspiration and Biomimetics*, *17*, 026012. doi: 10.1088/1748-3190/ac3f7f

Lopes, P., You, S., Cheng, L.-P., Marwecki, S., & Baudisch, P. (2017). Providing haptics to walls; heavy objects in virtual reality by means of electrical muscle stimulation. In *Proceedings of the 2017 chi conference on human factors in computing systems* (pp. 1471–1482). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/3025453.3025600 doi: 10.1145/3025453.3025600

Lougiakis, C., Katifori, A., Roussou, M., & Ioannidis, I.-P. (2020). Effects of virtual hand representation on interaction and embodiment in hmd-based virtual environments using controllers. In *2020 ieee conference on virtual reality and 3d user interfaces (vr)* (p. 510-518). doi: 10.1109/VR46266.2020.00072

Lugrin, J.-L., Latt, J., & Latoschik, M. (2015, 03). Avatar anthropomorphism and illusion of body ownership in vr. In *2015 ieee virtual reality (vr)* (p. 229-230). doi: 10.1109/VR.2015.7223379

Lukos, J., Snider, J., Hernandez, M., Tunik, E., Hillyard, S., & Poizner, H. (2013, 09). Parkinson's disease patients show impaired corrective grasp control and eye-hand coupling when reaching to grasp virtual objects. *Neuroscience*, *254*. doi: 10.1016/j.neuroscience.2013.09.026

Lyons, D. (1985). A simple set of grasps for a dextrous hand. In *Robotics and automation. proceedings. 1985 ieee international conference* (Vol. 2, p. 588 - 593).

Ma, S., Varley, M., Shark, L., & Richards, J. (2010). Emg biofeedback based vr system for hand rotation and grasping rehabilitation. In *2010 14th international conference information visualisation* (p. 479-484). doi: 10.1109/IV.2010.73

MacKenzie, C., & Iberall, T. (1994). Chapter 2. prehension. In C. L. Mackenzie & T. Iberall (Eds.), *The grasping hand* (Vol. 104, p. 15-46). North-Holland. Retrieved from https://www.sciencedirect.com/

science/article/pii/S0166411508615745 doi: https://doi.org/10.1016/S0166-4115(08)61574-5

Maldonado, J., & Zetzsche, C. (2021). Object manipulations in vr show task- and object-dependent modulation of motor patterns. In *Proceedings of the 27th acm symposium on virtual reality software and technology*. New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3489849.3489858 doi: 10.1145/3489849.3489858

Maranan, D. S., Fdili Alaoui, S., Schiphorst, T., Pasquier, P., Subyen, P., & Bartram, L. (2014, 04). Designing for movement: Evaluating computational models using lma effort qualities. In (p. 991-1000). doi: 10.1145/2556288.2557251

Martell, C. (2002, 01). Form: an extensible, kinematically-based gesture annotation scheme..

Martinet, A., Casiez, G., & Grisoni, L. (2010, 01). The effect of dof separation in 3d manipulation tasks with multi-touch displays. *Proceedings of VRST'10, the 17th ACM Symposium on Virtual Reality Software and Technology*. doi: 10.1145/1889863.1889888

Martirosov, S., Bures, M., & Zítka, T. (2022, 03). Cyber sickness in low-immersive, semi-immersive, and fully immersive virtual reality. *Virtual Reality*, *26*, 1-18. doi: 10.1007/s10055-021-00507-4

Masnadi, S., Pfeil, K. P., Sera-Josef, J.-V. T., & LaViola, J. J. (2021). Field of view effect on distance perception in virtual reality. In *2021 ieee conference on virtual reality and 3d user interfaces abstracts and workshops (vrw)* (p. 542-543). doi: 10.1109/VRW52623.2021.00153

Mason, M. T., & Salisbury, J. K. (1985). *Robot hands and the mechanics of manipulation*. Cambridge, MA, USA: MIT Press.

Matheus, K., & Dollar, A. (2010, 11). Benchmarking grasping and manipulation: Properties of the objects of daily living. In (p. 5020 - 5027). doi: 10.1109/IROS.2010.5649517

Mayer, E., Kriszun, K., Merz, L., Radon, K., Garrido, M. A., & Kranzlmüller, D. A. (2021). Designing an educational virtual reality application to learn ergonomics in a work place. *ACM International Conference on Interactive Media Experiences*.

McAweeney, E., Zhang, H., & Nebeling, M. (2018). User-driven design principles for gesture representations. In *Proceedings of the 2018 chi conference on human factors in computing systems* (p. 1-13). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3173574.3174121 doi: 10.1145/3173574.3174121

McBride, E. D. (1942). Disability evaluation. *The American Journal of Surgery, 3rd Edition*. doi: 10.1097/BLO.0b013e31803dcab1

Microsoft. (2021). *Get started with wpf.* Retrieved 2021-10-02, from https://docs.microsoft.com/en-us/visualstudio/designers/getting-started-with-wpf

Moehring, M., & Froehlich, B. (2011, March). Effective manipulation of virtual objects within arm's reach. In *2011 ieee virtual reality conference* (p. 131-138). doi: 10.1109/VR.2011.5759451

Monaco, S., Sedda, A., Cavina-Pratesi, C., & Culham, J. (2014, 11). Neural correlates of object size and object location during grasping actions. *European Journal of Neuroscience*, *41*. doi: 10.1111/ejn.12786

Morris, M. R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., schraefel, m. c., & Wobbrock, J. O. (2014, may). Reducing legacy bias in gesture elicitation studies. *Interactions*, *21*(3), 40-45. Retrieved from https://doi.org/10.1145/2591689 doi: 10.1145/2591689

Morris, M. R., Wobbrock, J. O., & Wilson, A. D. (2010). Understanding users' preferences for surface gestures. In *Proceedings of graphics interface 2010* (p. 261-268). CAN: Canadian Information Processing Society.

Motejlek, J., & Alpay, E. (2021). Taxonomy of virtual and augmented reality applications in education. *IEEE Transactions on Learning Technologies*, *14*(3), 415-429. doi: 10.1109/TLT.2021.3092964

Mousavi Hondori, H., Khademi, M., Dodakian, L., Cramer, S., & Lopes, C. (2013, 02). A spatial augmented reality rehab system for post-stroke hand rehabilitation. *Studies in health technology and informatics*, *184*, 279-85. doi: 10.3233/978-1-61499-209-7-279

Mu, S., & Sourin, A. (2021). Virtual assembling using hand tracking with leap motion controller. In *2021 international conference on cyberworlds (cw)* (p. 121-124). doi: 10.1109/CW52790.2021.00026

Muhanna, M. A. (2015). Virtual reality and the cave: Taxonomy, interaction challenges and research directions. *Journal of King Saud University - Computer and Information Sciences*, *27*(3), 344-361. Retrieved from https://www.sciencedirect.com/science/article/pii/S1319157815000439 doi: https://doi.org/10.1016/j.jksuci.2014.03.023

Multon, F., & Olivier, A.-H. (2013). Biomechanics of walking in real world: Naturalness we wish to reach in virtual reality. In F. Steinicke, Y. Visell, J. Campos, & A. Lécuyer (Eds.), *Human walking in virtual environments: Perception, technology, and applications* (pp. 55–77). New York, NY: Springer New York. Retrieved from https://doi.org/10.1007/978-1-4419-8432-6_3 doi: 10.1007/978-1-4419-8432-6_3

Murcia-López, M., & Steed, A. (2018). A comparison of virtual and physical training transfer of bimanual assembly tasks. *IEEE Transactions on*

Visualization and Computer Graphics, *24*(4), 1574-1583. doi: 10.1109/TVCG.2018.2793638

Murgia, A., & Sharkey, P. M. (2009, Jan.). Estimation of distances in virtual environments using size constancy. *International Journal of Virtual Reality*, *8*, 67-74. Retrieved from https://ijvr.eu/article/view/2714 doi: 10.20870/IJVR.2009.8.1.2714

Muse, D., Weber, C., & Wermter, S. (2006, 09). Robot docking based on omnidirectional vision and reinforcement learning. *Knowledge-Based Systems*, *19*, 324-332. doi: 10.1016/j.knosys.2005.11.018

Nabiyouni, M., Saktheeswaran, A., Bowman, D. A., & Karanth, A. (2015, March). Comparing the performance of natural, semi-natural, and non-natural locomotion techniques in virtual reality. In *2015 ieee symposium on 3d user interfaces (3dui)* (p. 3-10). doi: 10.1109/3DUI.2015.7131717

Napier, J. R. (1956, 12). The prehensile movements of the human hand. *The Journal of bone and joint surgery. British volume*, *38-B*, 902-13. doi: 10.1302/0301-620X.38B4.902

Nasim, K., & Kim, Y. J. (2016). Physics-based interactive virtual grasping. In *Proceedings of hci korea* (p. 114-120). Seoul, KOR: Hanbit Media, Inc. Retrieved from https://doi.org/10.17210/hcik.2016.01.114 doi: 10.17210/hcik.2016.01.114

Nayer, Z. H., Murdock, B., Dharia, I. P., & Belyea, D. A. (2020). Predictive and construct validity of virtual reality cataract surgery simulators. *Journal of Cataract Refractive Surgery*, *46*(6), 907–912.

Nebeling, M., Huber, A., Ott, D., & Norrie, M. C. (2014). Web on the wall reloaded: Implementation, replication and refinement of user-defined interaction sets. In *Proceedings of the ninth acm international conference on interactive tabletops and surfaces* (p. 15-24). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2669485.2669497 doi: 10.1145/2669485.2669497

Nichols, A., & Maner, J. (2008, 05). The good-subject effect: Investigating participant demand characteristics. *The Journal of General Psychology*, *135*, 151-65. doi: 10.3200/GENP.135.2.151-166

Nickerson, R., Varshney, U., & Muntermann, J. (2013, 05). A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, *22*. doi: 10.1057/ejis.2012.26

Nickerson, R., Varshney, U., Muntermann, J., & Isaac, H. (2007, 01). Towards a taxonomy of mobile applications. In *Amcis 2007 proceedings* (p. 338).

Nielsen, M., Stoerring, M., Moeslund, T., & Granum, E. (2003, 01). A procedure for developing intuitive and ergonomic gesture interfaces for hci. In (p. 409-420). doi: 10.1007/978-3-540-24598-8_38

Nieuwenhuizen, C. K. (2015). Quantitative analysis of computer interaction movements. In *Phd thesis*. Eindhoven University of Technology.

Nowak, K., & Biocca, F. (2003, 10). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence Teleoperators and Virtual Environments*, *12*, 481-494. doi: 10.1162/105474603322761289

Nunez, O., Zenner, A., Steinicke, F., Daiber, F., & Krüger, A. (2022, 06). Holitouch: Conveying holistic touch illusions by combining pseudo-haptics with tactile and proprioceptive feedback during virtual interaction with 3duis. *Frontiers in Virtual Reality*, *3*. doi: 10.3389/frvir.2022.879845

Oh, U., & Findlater, L. (2013). The challenges and potential of end-user gesture customization. In *Proceedings of the sigchi conference on human factors in computing systems* (p. 1129-1138). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/

2470654.2466145 doi: 10.1145/2470654.2466145

Ohwovoriole, M. S., & Roth, B. (1981, 10). An Extension of Screw Theory. *Journal of Mechanical Design*, *103*(4), 725-735. Retrieved from https://doi.org/10.1115/1.3254979 doi: 10.1115/1.3254979

Omair, B., & Alturki, A. (2020). An improved method for taxonomy development in information systems. *International Journal of Advanced Computer Science and Applications*, *11*.

Opris, D., Pintea, S., GarcÃa-Palacios, A., Botella, C., SzamoskÃ¶zi, Å., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: a quantitative meta-analysis. *Depression and Anxiety*, *29*(2), 85-93.

Ortega, F. R., Tarre, K., Kress, M., Williams, A. S., Barreto, A. B., & Rishe, N. D. (2019). Selection and manipulation whole-body gesture elicitation study in virtual reality. In *2019 ieee conference on virtual reality and 3d user interfaces (vr)* (p. 1110-1111). doi: 10.1109/VR.2019.8798182

Osso VR virtual reality surgical training platform. (n.d.). https://www.ossovr.com/. (Accessed: 2022-10-18)

Oudah, M., Al-Naji, A. A., & Chahl, J. (2020, 07). Hand gesture recognition based on computer vision: A review of techniques. *Journal of Imaging*, *6*, 73. doi: 10.3390/jimaging6080073

Paes, D., & Irizarry, J. (2018). A usability study of an immersive virtual reality platform for building design review: Considerations on human factors and user interface. In *Construction research congress 2018* (p. 419-428). Retrieved from https://ascelibrary.org/doi/abs/10.1061/9780784481264.041 doi: 10.1061/9780784481264.041

Papadimitriou, D. G. (2019). User experience in virtual reality, conducting an evaluation on multiple characteristics of a virtual reality experience. In *Human factors the journal of the human factors and ergonomics society*

(p. 964-982).

Park, H., Faghihi, N., Dixit, M., Vaid, J., & McNamara, A. (2021, 12). Judgments of object size and distance across different virtual reality environments: A preliminary study. *Applied Sciences*, *11*, 11510. doi: 10.3390/app112311510

Park, S.-M., & Kim, Y.-G. (2022). A metaverse: Taxonomy, components, applications, and open challenges. *IEEE Access*, *10*, 4209-4251. doi: 10.1109/ACCESS.2021.3140175

Patkin, M. (1981). Australian and New Zeeland Journal of Obstetrics and Gynaecology. *Ergonomics in Microsurgery*, *1*, 14-136.

Peillard, E., Thebaud, T., Normand, J.-M., Argelaguet, F., Moreau, G., & Lecuyer, A. (2019). Virtual objects look farther on the sides: The anisotropy of distance perception in virtual reality. In *2019 ieee conference on virtual reality and 3d user interfaces (vr)* (p. 227-236). doi: 10.1109/VR.2019.8797826

Pérez-Quiñones, M. A., Capra, R. G., & Shao, Z. (2003). The ears have it: A task by information structure taxonomy for voice access to web pages. In *Interact* (p. 856-859). IFIP TC13 International Conference on Human-Computer Interaction.

Pfeuffer, K., Mayer, B., Mardanbegi, D., & Gellersen, H. (2017). Gaze + pinch interaction in virtual reality. In *Proceedings of the 5th symposium on spatial user interaction* (p. 99-108). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3131277.3132180 doi: 10.1145/3131277.3132180

Pham, T., Vermeulen, J., Tang, A., & MacDonald Vermeulen, L. (2018). Scale impacts elicited gestures for manipulating holograms: Implications for ar gesture design. In *Proceedings of the 2018 designing interactive systems* *conference* (p. 227-240). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3196709.3196719 doi: 10.1145/3196709.3196719

Pillai, J. S., Ismail, A., & Charles, H. P. (2017). Grammar of vr storytelling: Visual cues. In *Proceedings of the virtual reality international conference - laval virtual 2017* (pp. 7:1–7:4). doi: 10.1145/3110292.3110300

Piumsomboon, T., Clark, A., Billinghurst, M., & Cockburn, A. (2013). User-defined gestures for augmented reality. In *Chi'13 extended abstracts on human factors in computing systems* (pp. 955–960).

Piumsomboon, T., Lee, G., Lindeman, R. W., & Billinghurst, M. (2017). Exploring natural eye-gaze-based interaction for immersive virtual reality. In *2017 ieee symposium on 3d user interfaces (3dui)* (p. 36-39). doi: 10.1109/3DUI.2017.7893315

Plumert, J. M., Kearney, J. K., Cremer, J. F., & Recker, K. (2005, jul). Distance perception in real and virtual environments. *ACM Trans. Appl. Percept.*, *2*(3), 216-233. Retrieved from https://doi.org/10.1145/1077399.1077402 doi: 10.1145/1077399.1077402

Pollard, N., & Lozano-Perez, T. (1990). Grasp stability and feasibility for an arm with an articulated hand. In *Proceedings., ieee international conference on robotics and automation* (p. 1581-1585 vol.3). doi: 10.1109/ROBOT.1990.126234

Prachyabrued, M., & Borst, C. W. (2014). Visual feedback for virtual grasping. In *2014 ieee symposium on 3d user interfaces (3dui)* (p. 19-26). doi: 10.1109/3DUI.2014.6798835

Prange, A., Barz, M., & Sonntag, D. (2018). *A categorisation and implementation of digital pen features for behaviour characterisation.*

Price, A. H., Kuttolamadom, M., & Obeidat, S. (2019). Using virtual reality welding to improve manufacturing process education.

Pusch, A., Martin, O., & Coquillart, S. (2008). Hemp-hand-displacement-based pseudo-haptics: A study of a force field application. *2008 IEEE Symposium on 3D User Interfaces*, 59-66.

Pusch, A., Martin, O., & Coquillart, S. (2009). Hemp-hand displacement-based pseudo-haptics: A study of a force field application and a behavioural analysis. *International Journal of Human-Computer Studies*, *67*(3), 256 - 268. Retrieved from http://www.sciencedirect.com/science/article/pii/S1071581908001328 (Current trends in 3D user interface research) doi: https://doi.org/10.1016/j.ijhcs.2008.09.015

Ragan, E., Bowman, D., Kopper, R., Stinson, C., Scerbo, S., & McMahan, R. (2015, 07). Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. *IEEE Transactions on Visualization and Computer Graphics*, *21*, 1-1. doi: 10.1109/TVCG.2015.2403312

Ranasinghe, N., Jain, P., Karwita, S., Tolley, D., & Do, E. Y.-L. (2017). Ambiotherm: Enhancing sense of presence in virtual reality by simulating real-world environmental conditions. In *Proceedings of the 2017 chi conference on human factors in computing systems* (p. 1731-1742). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3025453.3025723 doi: 10.1145/3025453.3025723

Ranasinghe, N., Jain, P., Thi Ngoc Tram, N., Koh, K. C. R., Tolley, D., Karwita, S., . . . Do, E. Y.-L. (2018). Season traveller: Multisensory narration for enhancing the virtual reality experience. In *Proceedings of the 2018 chi conference on human factors in computing systems* (p. 1-13). New

York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3173574.3174151 doi: 10.1145/3173574.3174151

Redmond, B., Aina, R., Gorti, T., & Hannaford, B. (2010). Haptic characteristics of some activities of daily living. In *Proceedings of the 2010 ieee haptics symposium* (pp. 71–76). IEEE Computer Society. doi: 10.1109/HAPTIC.2010.5444674

Reyes-Guzman, A., Dimbwadyo Terrer, I., Trincado-Alonso, F., Aznar, M., Alcubilla, C., Perez-Nombela, S., … Gil-Agudo, A. (2015, 07). Upper limb rehabilitation after spinal cord injury: a treatment based on a data glove and an immersive virtual reality environment. *Disabil Rehabil Assist Techno*.

Rietzler, M., Geiselhart, F., Gugenheimer, J., & Rukzio, E. (2018). Breaking the tracking: Enabling weight perception using perceivable tracking offsets. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 128:1–128:12). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/3173574.3173702 doi: 10.1145/3173574.3173702

Rodriguez, I. B., & Marquardt, N. (2017). Gesture elicitation study on how to opt-in and opt-out from interactions with public displays. In (p. 32-41). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3132272.3134118 doi: 10.1145/3132272.3134118

Rogers, K., Funke, J., Frommel, J., Stamm, S., & Weber, M. (2019). Exploring interaction fidelity in virtual reality: Object manipulation and whole-body movements. In *Proceedings of the 2019 chi conference on human factors in computing systems* (p. 414:1-414:14). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/3290605

.3300644 doi: 10.1145/3290605.3300644

Rosas, P., Wichmann, F., & Wagemans, J. (2007, 04). Texture and object motion in slant discrimination: Failure of reliability-based weighting of cues may be evidence for strong fusion. *Journal of vision*, *7*, 3. doi: 10.1167/7.6.3

Sahbi, H., Audibert, J.-Y., & Keriven, R. (2011). Context-dependent kernels for object classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*(4), 699-708. doi: 10.1109/TPAMI.2010.198

Schlesinger, G. (1919). Der mechanische aufbau der künstlichen glieder [the mechanical structure of artificial limbs]. M. Borchardt et al. (Eds.), Ersatznlieder und Arbeitshilfen fur Kriensbeschadiete und Unfallverletzte.

Schneider, T. W., & Balci, O. (2006). Vtquest: A voice-based multimodal web-based software system for maps and directions. In *Proceedings of the 44th annual southeast regional conference* (p. 300-305). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1185448.1185517 doi: 10.1145/1185448.1185517

Schwarz, R. J., & Taylor, C. (1955). The anatomy and mechanics of the human hand. *Artificial limbs*, *2*(2), 22–35.

Schwind, V., Knierim, P., Haas, N., & Henze, N. (2019). Using presence questionnaires in virtual reality. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 360:1–360:12). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/3290605.3300590 doi: 10.1145/3290605.3300590

Schwind, V., Knierim, P., Tasci, C., Franczak, P., Haas, N., & Henze, N. (2017). "these are not my hands!": Effect of gender on the perception of avatar hands in virtual reality. In *Proceedings of the 2017 chi conference on human factors in computing systems* (pp. 1577–1582). New York, NY, USA:

ACM. Retrieved from http://doi.acm.org/10.1145/3025453.3025602 doi: 10.1145/3025453.3025602

Scoditti, A., Blanch, R., & Coutaz, J. (2011). A novel taxonomy for gestural interaction techniques based on accelerometers. In *Proceedings of the 16th international conference on intelligent user interfaces* (p. 63-72). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1943403.1943414 doi: 10.1145/1943403.1943414

Seneler, C., Daim, T., & Basoglu, N. (2010, 01). Exploring user adoption of online services: A prototype driven framework for service design and development. *Journal of Entreprise Information Management*, *23*, 417-438.

Seneler, C. O., Basoglu, N., & Daim, T. U. (2008). A taxonomy for technology adoption: A human computer interaction perspective. In *Picmet '08 - 2008 portland international conference on management of engineering technology* (p. 2208-2219). doi: 10.1109/PICMET.2008.4599843

Sensorimotor control of grasping: Physiology and pathophysiology. (2009). Cambridge University Press. doi: 10.1017/CBO9780511581267

Serino, A., Alsmith, A., Costantini, M., Mandrigin, A., Tajadura-Jimenez, A., & Lopez, C. (2013). Bodily ownership and self-location: Components of bodily self-consciousness. *Consciousness and Cognition*, *22*(4), 1239 - 1252. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053810013001268 doi: https://doi.org/10.1016/j.concog.2013.08.013

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591–611.

Sharma, A., Roo, J. S., & Steimle, J. (2019). Grasping microgestures: Eliciting single-hand microgestures for handheld objects. In *Proceedings of the*

2019 chi conference on human factors in computing systems (p. 1-13). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3290605.3300632 doi: 10.1145/3290605.3300632

Shaw, E., Roper, T., Nilsson, T., Lawson, G., Cobb, S. V., & Miller, D. (2019). The heat is on: Exploring user behaviour in a multisensory virtual environment for fire evacuation. In *Proceedings of the 2019 chi conference on human factors in computing systems* (p. 1-13). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3290605.3300856 doi: 10.1145/3290605.3300856

Shen, W. (2021). Natural interaction technology in virtual reality. In *2021 international symposium on artificial intelligence and its application on media (isaiam)* (p. 1-4). doi: 10.1109/ISAIAM53259.2021.00008

Silva, E. S., Abreu, J., de Almeida, J. H. P., Teichrieb, V., & Ramalho, G. (2013). A preliminary evaluation of the leap motion sensor as controller of new digital musical instruments..

Simonsen, J. (2012). (T. R. Jesper Simonsen, Ed.).

Singh, A., Sha, J., Narayan, K. S., Achim, T., & Abbeel, P. (2014). Bigbird: A large-scale 3d database of object instances. *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 509-516.

Slocum, D. B., & Pratt, D. R. (1946). Disability evaluation for the hand. *JBJS*, *28*(3), 491–495.

Soh, J., Choi, Y., Park, Y., & Yang, H. S. (2013). User-friendly 3d object manipulation gesture using kinect. In *Proceedings of the 12th acm siggraph international conference on virtual-reality continuum and its applications in industry* (p. 231-234). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/

2534329.2534338 doi: 10.1145/2534329.2534338

Sollerman, C., & Ejeskär, A. (1995). Sollerman hand function test. a standardised method and its use in tetraplegic patients. *Scandinavian journal of plastic and reconstructive surgery and hand surgery*, *292*, 167-76.

Somers, D., & McNally, R. (2010, 10). Kinesthetic visual capture induced by apparent motion. *Journal of Vision - J VISION*, *3*, 35-35. doi: 10.1167/3.9.35

Song, J., Qin, S., & Zhang, P. (2016). Chinese text categorization based on deep belief networks. In *2016 ieee/acis 15th international conference on computer and information science (icis)* (p. 1-5). doi: 10.1109/ICIS.2016.7550914

Speicher, M., Lewis, K., & Nebeling, M. (2021, May). Designers, the stage is yours! medium-fidelity prototyping of augmented and virtual reality interfaces with 360theater. *Proc. ACM Hum.-Comput. Interact.*, *5*(EICS). Retrieved from https://doi.org/10.1145/3461727 doi: 10.1145/3461727

Spittle, B., Frutos-Pascual, M., Creed, C., & Williams, I. (2022). A review of interaction techniques for immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, 1-1. doi: 10.1109/TVCG.2022.3174805

Suhail, M., Sargunam, S. P., Han, D. T., & Ragan, E. D. (2017). Redirected reach in virtual reality: Enabling natural hand interaction at multiple virtual locations with passive haptics. In *2017 ieee symposium on 3d user interfaces (3dui)* (p. 245-246). doi: 10.1109/3DUI.2017.7893363

Sun, S. (2011, 12). Meta-analysis of cohen's kappa. *Health Services and Outcomes Research Methodology*, *11*. doi: 10.1007/s10742-011-0077-3

Sun, X., Liu, H., Tian, Y., Wu, G., & Gao, Y. (2020). Team effectiveness evaluation and virtual reality scenario mapping model for helicopter emergency rescue. *Chinese Journal of Aeronautics*. Retrieved from http://www.sciencedirect.com/science/article/pii/S1000936120302211 doi: https://doi.org/10.1016/j.cja.2020.06.003

Supuk, T., Bajd, T., & Kurillo, G. (2011, 05). Assessment of reach-to-grasp trajectories toward stationary objects. *Clinical biomechanics (Bristol, Avon)*, *26*, 811-8.

Suznjevic, M., Mandurov, M., & Matijasevic, M. (2017). Performance and qoe assessment of htc vive and oculus rift for pick-and-place tasks in vr. In *2017 ninth international conference on quality of multimedia experience (qomex)* (p. 1-3). doi: 10.1109/QoMEX.2017.7965679

Szabo, T., & Domokos, G. (2010, 03). A new classification system for pebble and crystal shapes based on static equilibrium points. *Central European Geology*, *53*, 1-19. doi: 10.1556/CEuGeol.53.2010.1.1

Tan, Y., Niu, C., & Zhang, J. (2020). Head-mounted, display-based immersive virtual reality marine-engine training system: A fully immersive and interactive virtual reality environment. *IEEE Systems, Man, and Cybernetics Magazine*, *6*(1), 46-51. doi: 10.1109/MSMC.2019.2948654

Tanjung, K., Farhan, L., Siregar, B., Panjaitan, S., & Fahmi, F. (2020, 05). The use of virtual reality controllers and comparison between vive, leap motion and senso gloves applied in the anatomy learning system. *Journal of Physics: Conference Series*, *1542*, 012026. doi: 10.1088/1742-6596/1542/1/012026

Tayag, M., Ronie, B., & Marvin, R. (2021). Leap motion controller enabled simulations of personal computer assembly programs. In *2021 the 4th international conference on software engineering and information management* (p. 17-21). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3451471.3451474 doi: 10.1145/3451471.3451474

Tian, H., Wang, C., Manocha, D., & Zhang, X. (2019). Realtime hand-object interaction using learned grasp space for virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, *25*(8), 2623-2635. doi: 10.1109/TVCG.2018.2849381

T. Iberall, D., Lyons. (1984). Towards perceptual robotics. *IEEE Intl Conf Systems*.

Tinguy, X. d., Pacchierotti, C., Emily, M., Chevalier, M., Guignardat, A., Guilllaudeux, M., ... Marchal, M. (2019). How different tangible and virtual objects can be while still feeling the same? In *2019 ieee world haptics conference (whc)* (p. 580-585). doi: 10.1109/WHC.2019.8816164

Tsoupikova, D., Stoykov, N. S., Corrigan, M., Thielbar, K., Vick, R., Li, Y., ... Kamper, D. G. (2014). Virtual immersion for post-stroke hand rehabilitation therapy. *Annals of Biomedical Engineering*, *43*, 467-477.

Tung, Y.-C., Hsu, C.-Y., Wang, H.-Y., Chyou, S., Lin, J.-W., Wu, P.-J., ... Chen, M. Y. (2015). User-defined game input for smart glasses in public space. In *Proceedings of the 33rd annual acm conference on human factors in computing systems* (p. 3327-3336). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2702123.2702214 doi: 10.1145/2702123.2702214

Ullmann, T., & Sauer, J. (2000). Intuitive virtual grasping for non haptic environments. In *Proceedings the eighth pacific conference on computer graphics and applications* (p. 373-457). doi: 10.1109/PCCGA.2000.883961

Ultraleap, L. M. (2021). *Leap motion controller specs - cnet.* Retrieved 2021-10-30, from https://www.cnet.com/products/leap-motion-controller/specs/

Valentini, P. P. (2018). Natural interface for interactive virtual assembly in augmented reality using leap motion controller. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, *12*, 1157-1165.

Verschoor, M., Lobo, D., & Otaduy, M. A. (2018). Soft hand simulation for smooth and robust natural interaction. In *2018 ieee conference on virtual reality and 3d user interfaces (vr)* (p. 183-190). doi: 10.1109/VR.2018.8447555

Viau, A., Feldman, A., McFadyen, B., & Levin, M. (2005, 01). Reaching in reality and virtual reality: A comparison of movement kinematics in healthy subjects and in adults with hemiparesis. *Journal of neuroengineering and rehabilitation*, *1*, 11. doi: 10.1186/1743-0003-1-11

Vigier, T., Moreau, G., & Siret, D. (2015, March). From visual cues to climate perception in virtual urban environments. In *2015 ieee virtual reality (vr)* (p. 305-306). doi: 10.1109/VR.2015.7223417

Villarreal-Narvaez, S., Vanderdonckt, J., Vatavu, R.-D., & Wobbrock, J. O. (2020). A systematic review of gesture elicitation studies: What can we learn from 216 studies? In *Proceedings of the 2020 acm designing interactive systems conference* (p. 855-872). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3357236.3395511 doi: 10.1145/3357236.3395511

Vogel, D., & Balakrishnan, R. (2005). Distant freehand pointing and clicking on very large, high resolution displays. In *Proceedings of the 18th annual acm symposium on user interface software and technology* (p. 33-42). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1095034.1095041 doi: 10.1145/1095034.1095041

Vogel, D., Lubos, P., & Steinicke, F. (2018). Animationvr - interactive controller-based animating in virtual reality. In *2018 ieee conference on virtual reality* *and 3d user interfaces (vr)* (p. 1-1). doi: 10.1109/VR.2018.8446550

Voigt-Antons, J.-N., Kojic, T., Ali, D., & Moeller, S. (2020). Influence of hand tracking as a way of interaction in virtual reality on user experience. In *2020 twelfth international conference on quality of multimedia experience (qomex)* (p. 1-4). doi: 10.1109/QoMEX48832.2020.9123085

Vuibert, V., Stuerzlinger, W., & Cooperstock, J. R. (2015). Evaluation of docking task performance using mid-air interaction techniques. In *Proceedings of the 3rd acm symposium on spatial user interaction* (p. 44-52). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2788940.2788950 doi: 10.1145/2788940.2788950

Waldow, K., Misiak, M., Derichs, U., Clausen, O., & Fuhrmann, A. (2018). An evaluation of smartphone-based interaction in ar for constrained object manipulation. In *Proceedings of the 24th acm symposium on virtual reality software and technology* (p. 1-2). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3281505.3281608 doi: 10.1145/3281505.3281608

Wan, H., Luo, Y., Gao, S., & Peng, Q. (2004). Realistic virtual hand modeling with applications for virtual grasping. In *Proceedings of the 2004 acm siggraph international conference on virtual reality continuum and its applications in industry* (p. 81-87). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1044588.1044603 doi: 10.1145/1044588.1044603

Wang, Q., Kang, B., & Kristensson, P. O. (2021). Supporting physical and mental health rehabilitation at home with virtual reality headsets and force feedback gloves. In *2021 ieee conference on virtual reality and 3d user interfaces abstracts and workshops (vrw)* (p. 685-686). doi: 10.1109/

VRW52623.2021.00225

Wang, Y., Su, Z., Zhang, N., Liu, D., Xing, R., Luan, T. H., & Shen, X. S. (2022). A survey on metaverse: Fundamentals, security, and privacy. *ArXiv*, *abs/2203.02662*.

Ware, C., & Arsenault, R. (2004). Frames of reference in virtual object rotation. In *Proceedings of the 1st symposium on applied perception in graphics and visualization* (p. 135-141). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1012551.1012576 doi: 10.1145/1012551.1012576

Whitwell, R., Ganel, T., Byrne, C., & Goodale, M. (2015, 01). Real-time vision, tactile cues, and visual form agnosia in pantomimed grasping: Removing haptic feedback induces a switch from natural to pantomimelike grasps. *Front. Hum. Neurosci.*, *9*, 9-216.

Willemsen, P., Colton, M. B., Creem-Regehr, S. H., & Thompson, W. B. (2004). The effects of head-mounted display mechanics on distance judgments in virtual environments. In *Proceedings of the 1st symposium on applied perception in graphics and visualization* (p. 35-38). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1012551.1012558 doi: 10.1145/1012551.1012558

Williams, A. S., Garcia, J., & Ortega, F. (2020). Understanding multimodal user gesture and speech behavior for object manipulation in augmented reality using elicitation. *IEEE Transactions on Visualization and Computer Graphics*, *26*(12), 3479-3489. doi: 10.1109/TVCG.2020.3023566

Williams, B., Narasimham, G., Westerman, C., Rieser, J., & Bodenheimer, B. (2007, jul). Functional similarities in spatial representations between real and virtual environments. *ACM Trans. Appl. Percept.*, *4*(2), 12-es. Retrieved from https://doi.org/10.1145/1265957.1265961 doi: 10.1145/1265957.1265961

Wimmer, R. (2011). Grasp sensing for human-computer interaction. In *Proceedings of the fifth international conference on tangible, embedded, and embodied interaction* (pp. 221–228). New York, NY, USA: ACM. doi: 10.1145/1935701.1935745

Witmer, B. G., & Singer, M. J. (1998, June). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225-240. doi: 10.1162/105474698565686

Wittorf, M. L., & Jakobsen, M. R. (2016). Eliciting mid-air gestures for wall-display interaction. In *Proceedings of the 9th nordic conference on human-computer interaction* (p. 3).

Wobbrock, J. O., Aung, H. H., Rothrock, B., & Myers, B. A. (2005). Maximizing the guessability of symbolic input. In *Chi '05 extended abstracts on human factors in computing systems* (pp. 1869–1872). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1056808.1057043 doi: 10.1145/1056808.1057043

Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). User-defined gestures for surface computing. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 1083–1092). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1518701.1518866 doi: 10.1145/1518701.1518866

Wobbrock, J. O., Myers, B. A., & Kembel, J. A. (2003). Edgewrite: A stylus-based text entry method designed for high accuracy and stability of motion. In *Proceedings of the 16th annual acm symposium on user interface software and technology* (p. 61-70). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/

964696.964703 doi: 10.1145/964696.964703

Won, S., & Westland, S. (2017). Colour meaning and context. *Color Research & Application*, *42*(4), 450-459. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/col.22095 doi: 10.1002/col.22095

Wu, B., Ooi, T. L., & He, Z. J. (2004). Perceiving distance accurately by a directional process of integrating ground information. *Nature*, *428*, 73-77.

Wu, H., Luo, W., Pan, N., Nan, S., Deng, Y., Fu, S., & Yang, L. (2019, 12). Understanding freehand gestures: a study of freehand gestural interaction for immersive vr shopping applications. *Human-centric Computing and Information Sciences*, *9*, 43. doi: 10.1186/s13673-019-0204-7

Xu, W., Liang, H.-N., He, A., & Wang, Z. (2019). Pointing and selection methods for text entry in augmented reality head mounted displays. In *2019 ieee international symposium on mixed and augmented reality (ismar)* (p. 279-288). doi: 10.1109/ISMAR.2019.00026

Yan, Y., Yu, C., Ma, X., Yi, X., Sun, K., & Shi, Y. (2018). Virtualgrasp: Leveraging experience of interacting with physical objects to facilitate digital object retrieval. In *Proceedings of the 2018 chi conference on human factors in computing systems* (p. 1-13). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3173574.3173652

Yang, L., Chen, J., & Zhu, W. (2020, 04). Dynamic hand gesture recognition based on a leap motion controller and two-layer bidirectional recurrent neural network. *Sensors*, *20*, 2106. doi: 10.3390/s20072106

Yim, D., Loison, G. N., Fard, F. H., Chan, E., McAllister, A., & Maurer, F. (2016). Gesture-driven interactions on a virtual hologram in mixed reality. In *Proceedings of the 2016 acm companion on interactive surfaces and spaces*

(p. 55-61). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3009939.3009948 doi: 10.1145/3009939.3009948

Yu, D., Liang, H.-N., Zhang, T., & Xu, W. (2019). Depthmove: Hands-free interaction in virtual reality using head motions in the depth dimension. In *2019 ieee conference on virtual reality and 3d user interfaces (vr)* (p. 1257-1258). doi: 10.1109/VR.2019.8797916

Zahariev, M., & MacKenzie, C. (2007, 07). Grasping at 'thin air': Multimodal contact cues for reaching and grasping. *Experimental brain research. Experimentelle Hirnforschung. ExpÃ©rimentation cÃ©rÃ©brale*, *180*, 69-84. doi: 10.1007/s00221-006-0845-4

Zenner, A., & Krüger, A. (2017). Shifty: A weight-shifting dynamic passive haptic proxy to enhance object perception in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, *23*(4), 1285-1294. doi: 10.1109/TVCG.2017.2656978

Zhang, Z., Cao, B., Weng, D., Liu, Y., Wang, Y., & Huang, H. (2018, March). Evaluation of hand-based interaction for near-field mixed reality with optical see-through head-mounted displays. In *2018 ieee conference on virtual reality and 3d user interfaces (vr)* (p. 739-740). doi: 10.1109/VR.2018.8446129

Zheng, J. Z., Rosa, S. D. L., & Dollar, A. M. (2011). An investigation of grasp type and frequency in daily household and machine shop tasks. *2011 IEEE International Conference on Robotics and Automation*, 4169-4175.

Ziemer, C., Plumert, J., Cremer, J., & Kearney, J. (2006). Making distance judgments in real and virtual environments: Does order make a difference? In *Proceedings of the 3rd symposium on applied perception in graphics and visualization* (p. 153). New York, NY, USA: Association for

Computing Machinery. Retrieved from https://doi.org/10.1145/1140491.1140530 doi: 10.1145/1140491.1140530

Zingg, T. (1935). *Beitrag zur schotteranalyse* (Unpublished doctoral dissertation). ETH Zürich,Zürich, Switzerland.