Institute for Visualization and Interactive Systems

University of Stuttgart
Universitätsstraße 38
D–70569 Stuttgart

Masterarbeit

A Literature Review on Distant Object Selection Methods

Jan Leusmann

Course of Study:	Informatik
Examiner:	Prof. Dr. Michael Sedlmair
Supervisor:	Prof. Dr. Sven Mayer Katrin Angerbauer, M.Sc.
Commenced:	March 4, 2021
Completed:	September 3, 2021

Abstract

Nowadays, we have a large quantity of different Virtual and Augmented Reality devices with ever-improving technology. Although those devices are mainly still used for entertainment purposes, we also see new applications regarding data analysis, virtual meetings, and many more. We observe that these different applications provide environments, which are vastly different in terms of their target density and occlusion. Many application-specific interaction techniques were proposed over the last decades to solve the challenges coming from diverse environments. Users, who want to use multiple applications, often have to switch between various techniques for interaction and learn entirely new application-specific selection techniques, especially for distant target selection. We propose the need for a design space to classify selection techniques and streamline the development of an all-encompassing selection technique. For this purpose, we present a systematic literature review. We include 146 records using the PRISMA guidelines. Through an investigation of the literature, we extract ten impact factors to classify distant selection techniques. We present a design space with these factors clustered into three categories: input and output devices, the selection process, and the confirmation of the selection. Supported by this design space, the development of new selection techniques can be more effective in creating one technique, able to solve challenges from complex environments while still leading to little user fatigue and high levels of immersion.

Kurzfassung

Heutzutage existieren eine große Anzahl verschiedener Virtual- und Augmented-Reality-Hardware mit sich kontinuierlich verbessernder Technologie. Obwohl diese hauptsächlich zu Unterhaltungszwecken eingesetzt werden, gibt es auch neue Anwendungen für Datenanalysen, virtuelle Meetings und vieles mehr. Innerhalb dieser verschiedenen Anwendungen können Objekte unterschiedlich dicht aneinander platziert oder verdeckt sein. In den letzten Jahrzehnten wurden deshalb viele anwendungsspezifische Interaktionstechniken vorgeschlagen, um die Herausforderungen unterschiedlicher Umgebungen zu lösenBenutzer, die mehrere Anwendungen nutzen wollen, müssen oft zwischen unterschiedlichen Interaktionstechniken wechseln und völlig neue anwendungsspezifische Selektier-Methoden erlernen, insbesondere für die Selektion entfernter Ziele. Wir schlagen deshalb einen Design-Space vor, um Auswahltechniken zu klassifizieren und die Entwicklung einer allumfassenden Selektier-Methode zu vereinheitlichen. Zu diesem Zweck präsentieren wir eine systematische Literaturanalyse. Wir berücksichtigen 146 Veröffentlichungen unter Anwendung der PRISMA-Richtlinien. Aus dieser Literatur extrahieren wir zehn Einflussfaktoren, um entfernte Selektier-Methoden zu klassifizieren. Diese Arbeit stellt einen Design Space vor, in dem diese Faktoren in drei Kategorien unterteilt sind: Eingabe- und Ausgabegeräte, der Auswahlprozess und die Bestätigung der Auswahl. Mit Hilfe eines solchen Design Spaces soll die Entwicklung neuer Selektier-Methoden effektiver zu gestaltet werden. So können Techniken geschaffen werden, die in der Lage sind, Herausforderungen in komplexen Umgebungen zu lösen und zudem dem Nutzer eine möglichst wenig anstrengende und immersive Erfahrung bieten.

Contents

List of Figures

List of Tables

1 Introduction

Our primary way of interacting with digital environments is selection. On a PC, we need selection for, e.g., accessing files, choosing data points in graphs, following links on webpages, or using tools and interacting with characters in video games. Selection processes on digital devices are performed almost subconsciously, as mechanisms evolved to be as natural as grasping a door handle. Whereas on a computer, selection is still mostly done with a mouse, on modern smartphones the selection has evolved from pressing buttons to the more natural way of touch interaction. Digital selection is ubiquitous. Virtual Reality (VR) at this point is not a new technology anymore and went through the Gartner hype cycle multiple times already. The first VR-Head-Mounted-Displays (HMDs) were available since the late 1980s [153]. At the latest, with the release of multiple consumer-ready VR-HMDs like the Oculus Rift and HTC Vive in 2016, VR became something more consumers, and thus, also researchers found interest in. Most recently, the release of the Oculus Quest 2, a standalone VR-HMD, made it clear that many consumers are ready to buy this new form of digital entertainment system. In the first six months of its release, the Oculus Quest 2 “has outsold not just its predecessor but all of its predecessors combined”[1]. However, even though sales of VR glasses are on the rise, there are still many challenges researchers need to overcome.

One of these challenges is finding a selection method equivalent to the mouse on a PC for VR. Many selection methods have been proposed, and new ones are getting developed every year. Nevertheless, we found that there is currently no way for developers to compare them quickly. Thus, the user often has to learn new ways of interacting with a virtual environment. We propose that the main reason for this magnitude of selection techniques for VR comes from the technical evolution of VR. As hardware and software progressed in the last 20 years, multiple researchers tried to find an optimal solution for target selection in VR. Releases of devices like the Leapmotion[2] and the Oculus Quest 2, including hand tracking[3], made hand tracking more available than ever. Thus, many researchers tried to develop novel selection methods with hand tracking as their input modality [152, 181]. Eye-tracking technology also evolved in the last decade, which, again, brought up many studies using eye-tracking technologies as an input modality for VR [85, 119, 150, 151]. The most ubiquitous input method is still the 6-Degrees-of-Freedom (DoF) controller. However, here too, there are many different types of controllers and how they get used. All these different input modalities vary in accuracy depending on the quality of used technology. Thus, many researchers tried to solve the problem of accuracy on their own, bringing us to the current state of having to choose between many different selection techniques without a good way of comparing them. Furthermore, although there are some literature reviews regarding this topic [3, 70, 107, 183], we propose that the design space for VR selection is not clearly defined. This enabled researchers to

---

[1] https://www.bloomberg.com/news/videos/2021-03-29/facebook-reality-labs-vp-vr-will-transform-global-work-video

[2] https://www.ultraleap.com/product/leap-motion-controller/

[3] https://www.oculus.com/quest-2/

have a good way of classifying their created selection method into different categories and research underdeveloped areas. The scope of these literature reviews was not to group and classify methods, but compare single methods. Hence, our solution to this problem is to systematically review the literature on selection methods on a large scale to find factors to classify methods. Bergström et al. [18] did a similar survey but on the topic of how to evaluate object selection and manipulation studies. They gave researchers a structured overview of how they should design their study for a new selection method. Similar to their approach, we want to provide an overview of the current state of selection methods. We aim to guide future researchers by clarifying which ideas have already been researched to which extend and find gaps in research that have not been thoroughly investigated. Our goal is to propose a design space to classify current selection technologies and guide future work.

In this work we present a systematic literature on the topic of distant selection for 3D-environments. We followed the PRISMA guidelines [113], and collected 146 different papers on the topic of distant selection. Out of this pool of literature we extract impact factors, and classify selection techniques from the literature into descriptive classes, to then evaluate advantages and disadvantages of different methods. From this we propose a design space for distant selection techniques, with three main categories: input and output devices, selection process, and confirmation of the selection. Each category has multiple factors with fitting levels. This design space should help future researchers streamline the process of developing new selection methods with the goal of unifying 3D interaction for different applications. Furthermore, we discuss the two main challenges of designing distant selection techniques: dense & occluded environments and usability of selection techniques.

2 Related Work

This chapter provides an overview of the history of selection in VR, argues why there are many different interaction techniques, and summarizes earlier classification methods. Additionally, we explore the reality-virtuality continuum and current VR research topics to motivate reasons why we need a standardized selection method.

2.1 Selection in Virtual Reality

Selection is an indispensable means to interact with digital spaces, including virtual environments. In the early phases of VR, the main challenge was to bring the concept of a virtual environment through an HMD to life. Thus, the first interaction techniques only came in the subsequent technological advances of VR. In 1961, one year after the first stereoscopic HMD, patented by Moron Heilig [55], Headsight was developed, which included motion tracking to enable a change of the viewport [31]. These HMDs were far from VR, as we know it today, but could already display stereoscopic video material. And although the technology for an immersive virtual experience had yet to be invented, Ivan Sutherland, in his essay The Ultimate Display, already predicted a very close image to today's virtual environments [163]. Sutherland later created the first VR-HMD, The Sword of Damocles, which still did not include any interaction methods but was another step in the technological advancement of VR. In the following decades, many developers released multiple new VR-HMDs. Not only optical advances drove the development forward, but also the first input devices were invented, mainly being glove-based hand-tracking systems. The main challenge of these early systems was to translate existing knowledge from 2D selection tasks to a 3D space [91]. VR HMDs coming out in the following decades started to tackle how interaction in VR should work. VPL Research was the first company to sell VR headsets and gloves for input, which was a milestone for the consumer VR market [200]. Due to the technical means available at the time, the tracking accuracy of these input devices was below average. Thus, researchers tried to overcome these problems by building custom hardware and creating their own selection methods. Freeman and Weissman [42] designed one of the earliest freehand pointing systems.

Today, we have multiple tracking solutions for the same input method, varying in accuracy and feasibility. Tracking the hands in VR can be done via gloves (e.g., ManusGlove[1]), marker-based systems (e.g., Optitrack[2]), or camera based system (e.g., Leapmotion[3]). Most consumer VR-HMDs today come with a pair of 6-DoF controllers as their primary tool to interact with the virtual

---

[1] `https://www.manus-vr.com/`

[2] `https://optitrack.com/`

[3] `https://www.ultraleap.com/product/leap-motion-controller/`

environment[4]. Less expensive headsets, like Google Cardboard[5], still rely on head-tracking for selection; others included eye-tracking[6]. The challenge shifted away from "how to provide interaction in VR on a technical level" as we do not have a lack of sufficiently good hardware anymore. Today, we have to solve the question, "what method is best for selection in VR", as object selection in 3D spaces is one of the fundamental tasks [81]. To solve this question, we have to build and evaluate a foundation of techniques to create one consistent 3D distant target selection method. We can split 3D object selection techniques into two main categories: virtual hand [132] and virtual pointing [91, 112]. In the early days of VR, virtual environments tried to replicate the real world as much as possible. Thus, virtual hand techniques were the first intuitive idea to interact with virtual environments. Due to the possibility of interaction in a virtual environment going beyond the borders of the real environment, selection got expanded into distant selection, using the metaphor of pointing at objects. The most ubiquitous method to implement selection via pointing is raycasting [112], where a vector is built from an origin position and a direction. This ray can either be of infinite or finite length, but the first object hit by the ray is selected. Virtual hand techniques are, mainly, near target selection only. In this literature review, we focus on distant target selection; thus, the majority of work we will look at will be virtual pointing techniques.

Due to the large variety of hardware solutions, applications, and users' needs, a wide array of interaction techniques got invented over the years. Moreover, multiple classification methods got proposed over the last decades. Poupyrev and Ichikawa [134] in 1999 suggested to classify selection methods based on them being egocentric or exocentric. Additionally, Bowman et al. [22] in 2001 proposed to divide the interaction process into three subtasks: indication of object, confirmation of selection, and feedback. In 2013 Argelaguet and Andujar [3] performed a systematic literature review. They collected 32 different 3D selection methods and proposed classification guidelines as they found difficulties with the classification methods by Bowman et al. [22] and Poupyrev and Ichikawa [134]. Furthermore, they propose that selection methods should be classified depending on their "intrinsic characteristics, underlying selection techniques, and how the user controls it" [3]. Thus, the authors proposed the following factors to classify selection techniques: *selection tool* describes shape and properties of the virtual cursor, *tool control* covers technical characteristics of the input tool, *motor and visual space relationship* describes how physical positions are transferred to the virtual space, *disambiguation mechanism* defines the decision process of which target is selected with, e.g., heuristics, *selection trigger* specifies how a selection is confirmed, and *feedback* includes all aspects, visual, haptic, and auditory, happening during the process and after a performed selection. With these classes, they strive to label the critical aspects in differentiating selection methods. Weise et al. [183] proposed another set of classes intending to provide an extensive characterization of interaction techniques. They extracted 13 classes: *metaphor* describes if interaction is performed through grasping, pointing, or a hybrid solution, *task* divides interaction into selection, positioning, rotation, and scaling, *bimanual* states if both hands are used for interaction, *DoF* indicates the amount of input dimensions, *constraints* characterize restrictions by the virtual environment, *Control-Display-Ratio* describes if the input device uses a 1-to-1 mapping with regards to the virtual cursor, *spatial compliance* illustrates the consistency between the real world movement and virtual movement, *input device* describes which modalities are used by this interaction technique, *reference frame* is split up into egocentric and exocentric, *action space* describes the radius in which users

---

[4] `https://www.oculus.com/rift-s/`, `https://www.vive.com/de/`

[5] `https://arvr.google.com/cardboard/`

[6] `https://www.microsoft.com/de-de/hololens`

can interact with virtual objects, *disambiguation* tells if and which heuristics are applied in the selection process, *directness* divides interaction techniques into indirect and direct techniques, and *interaction fidelity* decouples the level of realism in interaction from immersion.

In addition to their classification factors, Argelaguet and Andujar [3] discussed usability factors influencing selection performance. They propose that a selection technique should enable users to perform fast and accurate selections while being easy to understand and not fatiguing. In order to achieve these goals, they additionally provide researchers with a set of challenges, which have to be solved, depending on the environment being dense or sparse. *Target geometry*, *object distance*, and *object density* describe that pointing at small, distant, or occluded targets can be challenging. With the factor *input and output devices*, they describe the challenge of having too many input and output devices, which can make it complicated to choose the best devices or make a technique available for every possible device. Additionally challenges are *user fatigue*, *user preference*, and lastly *application performance, latency, and noise*. To cope with noise from the input device or the user's hand trembling, either a band-pass filter or Kalman filter should be applied [78]. Two additional challenges for distant selection methods are Midas Touch [64] and the Heisenberg effect [21]. Midas Touch originates from the field of eye-tracking and defines that "everywhere you look, something is activated; you cannot look anywhere without issuing a command" [64]. Meaning that a trigger to confirm a selection is necessary, as otherwise every selectable object will always be directly selected when hovering over it. The Heisenberg effect describes the problem that the cursor's position will likely change when confirming a selection with a physical motion due to the user's hand tremble. This effect can be circumvented by, e.g., confirming selections with the hand which is not used to perform the pointing process or using dwell time as a trigger.

Solid performance and accuracy are integral elements for any selection method. To compare the performance of 2D selection methods, Fitts' Law is a widely accepted method [41]. Over the years, many variations of the original Fitts' Law have been proposed and used. Due to its ubiquitousness, many researchers have also used a Fitts' Law task to evaluate the performance of their selection techniques in virtual environments. And despite the high demand for a standardized performance metric, no performance metric for 3D selection tasks has yet prevailed. Triantafyllidis and Li [171] provide an overview of multiple different versions of Fitts' Law used in VR studies and address the difficulties of finding a unified 3D performance model due to the multitude of vastly different input devices. Bergström et al. [18] performed a systematic literature review on how to evaluate object selection in VR. They also found that there are yet no design guidelines in designing studies regarding selection in 3D spaces, and therefore, evaluating the performance of different studies is difficult. To tackle this and bring future selection methods to a point where evaluation across different studies can be done, they propose ten design guidelines to help future researchers deciding how to design their studies.

We propose that due to the high amount of different selection methods, there is a need for an overview of already investigated selection methods to identify gaps in this research topic. Previous literature reviews came to varying levels of fine granularity in classifying selection methods. Due to the ever-improving level of VR technology, many previous challenges to solve distant selection are no longer present. To create one universal distant selection method, we must identify current challenges and ways to solve them. This universal selection method should satisfy user needs on a performance level and be easy to use, not interrupting immersion, and not fatiguing the users. A suitable interaction method can even increase levels of perceived immersion. This work should be in line with other current works of identifying ways to compare selection methods from different studies and cover the whole space of 3D interaction.

2.2 Reality-Virtuality Contiuum

The term mixed reality (MR) got coined by Milgram and Kishino in 1994 [109]. They state that the real environment, augmented reality (AR), augmented virtuality, and virtual environment are all somewhere along the virtuality continuum. Current research widely accepts this continuum. HMDs allow users to transition from real worlds to virtual environments or augment reality. Interacting with the environment on the continuum is injektiv [103]. We can perform every type of interaction we can do in a real environment, in AR and VR. However, interaction in VR and AR can even expand the types of interaction we can do in the real world. In a computer-generated virtual environment, information about every object exists, thus, making it selectable or even manipulable. Although today primarily created with HMDs, virtual environments can also be perceived with a Cave Automatic Virtual Environment (CAVE) or stereo displays. AR can be perceived either through optical see-through or video see-through [140]. Interaction techniques for these environments all need to be three-dimensional. As distant selection is possible in every environment, except the real one, we can evaluate 3D selection techniques in all these environments.

Interacting with distant targets is not possible in the real world but is widely used in virtual and augmented environments. Although the gaming industry still has the largest market share in VR, businesses also start to see and use the advantages of virtual environments[7]. Research in the medical industry has started to use VR for topics like phobia treatment [36, 111], explored virtual rehabilitation training for patients with Parkinson's disease [87], and also virtual surgical training [62]. There is also work in how to augment actual operating rooms with hands-free interaction [53]. Furthermore, in other settings like collaborative interaction or data analysis tasks, distant selection can also significantly increase productivity and accessibility. Millais et al. [110] compared data visualizations in 2D environments and virtual environments. Users could freely navigate through the virtual environment to immersive themselves in the dataset. They found that users were more satisfied with their data explorations in VR. In these kinds of environments, the previously mentioned challenges of object density and occlusion, are ubiquitous and to enable smooth interaction with them, these challenges should be solved. Tadeja et al. [165] investigated the use of Parallel Coordinate plots in virtual environments. Interaction in their environment was done with a gamepad and did not solve the challenges coming from dense and occluded environments. They state that their results could have been better regarding user experience in VR with a more sophisticated selection method. Büschel et al. [23] proposed an MR application to enable collaborative interaction for spatio-temporal data, where the data could be viewed on a large screen and additionally manipulated in AR. Due to the 3D environment of the augmented environment, data trends and spatial aspects of the data could be found more easily. The benefits of 3D environments were also incorporated into different topics like teaching and virtual showcases.

---

[7] `https://www.grandviewresearch.com/industry-analysis/virtual-reality-vr-market`

Although such applications are great for users to have new possibilities in how they explore data, learn, and experience things, currently, they often have to relearn how to interact with these applications. The need for different selection methods for different applications shows that the current universal distant selection methods are insufficient in covering the vast array of challenges presented by different environments. If we want to make such applications accessible for every user, a standardized interaction method should be used, which can still manage to select occluded objects where object density is high.

2.3 Summary

Even though VR is already getting researched since the 1980s, the latest releases of accessible VR-HMDs made a massive spike in consumer and research interest. As the necessary technology for VR progressed, researchers started to propose different selection techniques for different input devices. With advances in input and output technologies the challenge of 3D-interaction shifted towards which selection techniques can solve different environmental challenges. We already have sophisticated interaction techniques for specific use cases, however no unified method covers the wide range of different environments and applications. Furthermore, specific input devices have prevailed for different environments from the reality-virtuality continuum. In augmented environments, hand tracking and eye gaze are mostly used as input modalities and controllers in virtual environments. Users who want to engage in different applications across the continuum often have to learn new selection methods for a specific application. This dramatically hinders their productivity, as they can not rely on their previous knowledge of how to interact with 3D environments. Although there are reasonable solutions for specific use cases, we propose the need for a design space to learn the underlying factors, classify selection techniques, and generalize selection in 3D spaces.

3 Systematic Literature Review

In this chapter, we will lay out our methodological process how we found suiting literature. For this, we followed the four phases of the PRISMA guidelines [113] (identification, screening, eligibility, inclusion) to create our dataset of existing work and ensure transparency and reproducibility of our process. This process is displayed in Figure 3.1. We aimed to categorize literature so that we could later identify commonalities and differences. In contrast to previous work [3, 183], we included every publication using a selection method and not only newly proposed techniques.

3.1 Identification

The first step was to identify keywords for our search query. We built upon the investigations from Argelaguet and Andujar [3] and included their compared literature into our dataset. Additionally, we investigated Google Scholar to situate ourselves within existing research. Through an investigation of this literature, we gained knowledge on possible design factors. As we situate our work in the space of VR, we first added equivalent terms to our search query. Within our first investigation of the literature on Google Scholar, we found several works on selection methods for large displays, which were also applicable in virtual environments as pointing towards the screen was part of the selection process [35, 66, 178]. Due to the similarities between VR and AR, we decided to also include *AR* as a term into our search query to gain novel selection method ideas from this related space. In contrast to the literature review by Monteiro et al. [115], who looked into a specific sub-topic of interaction in VR, we wanted to conclude our work based on a broader spectrum of selection methods for VR. Thus, we also included *selection*, *pointing*, and related terms into our search query. We then investigated the number of search results for different keyword combinations to find queries with a feasible but sufficient amount of literature. One of our metrics to decide the quality of the current keywords was to find as many papers we previously found on Google Scholar as possible. We finally decided to combine the following *method-keywords* with *environment-keywords*, which left only five papers from Google Scholar unfound. We performed one query search for each combination of the following method and environment keywords.

```
('pointing', 'selection', 'mid-air pointing', 'target-selection', 'target selection',
 'gesture')
AND
('virtual environment', 'virtual reality', 'vr' OR 'large screen', 'large display',
 'augmented reality')
```

In contrast to the work by Argelaguet and Andujar [3], who handpicked selection methods for their systematic review, we wanted to gather as much literature on selection methods as possible to extract design factors. Due to the technical advances of VR in the last five years and the accompanying rising interest of researchers, a large corpus of literature was published over the last years, which we mainly wanted to include. To tackle this problem systematically, we developed a Python-application

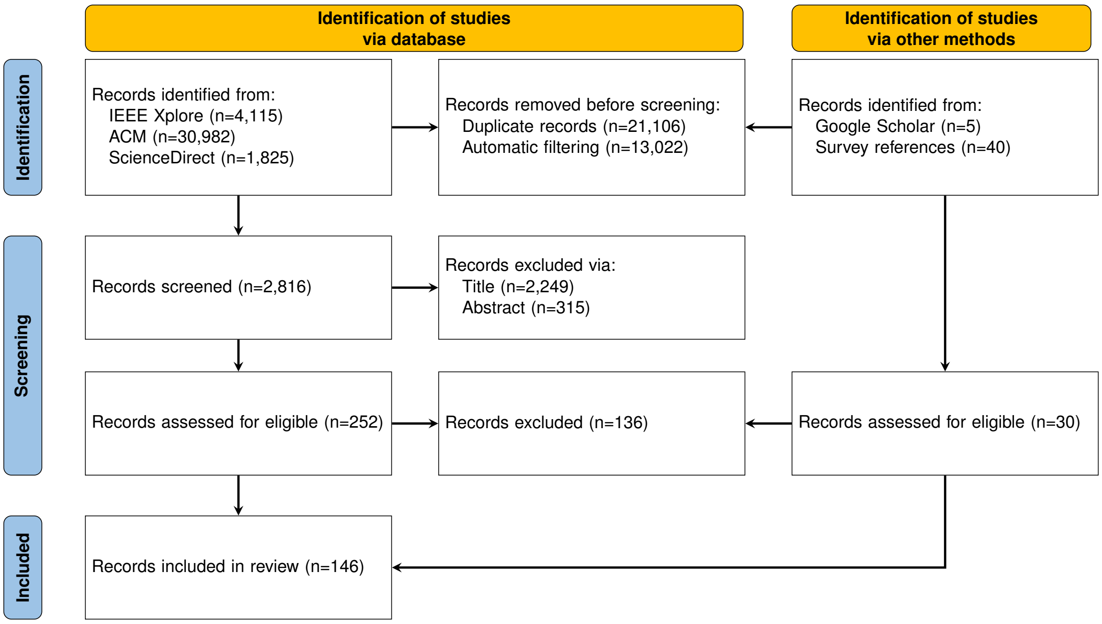


Figure 3.1: PRISMA flow diagram to show the process of our systematic literature analysis.

which crawled every entry fitting to this search query. We conducted the query search in July of 2021. We decided not to limit search results to a specific period and applied our search query with default settings on the libraries: IEEE, ACM, and ScienceDirect. In some cases, this resulted in more search results than the page number limit set by the database. Thus, we were not able to gather every paper fitting to that search query. To deal with this, we additionally applied our search query with the setting *title-search* to every library. With this, we hoped to identify at least the most fitting works for every search query.

In this step, we gathered 36,922 entries from the three libraries. We then removed duplicates which resulted in 15,838 entries. As this amount of entries was still far too high to screen systematically, we handpicked some of the entries to see if they were eligible and find universal reasons to exclude papers from our database systematically. We found that many papers were not on the topic of selection in VR but instead were on entirely different topics. We suspect that query words were part of words used in these papers. Thus, we filtered the resulting entries again by taking our query words and only including papers where the query words would appear in either the title or abstract. This resulted in 2,816 entries which we could then screen systematically.

3.2 Screening

In this part, we manually screened the resulting entries to chose the papers we wanted to include in our dataset. We did this in two phases: title-screening and abstract-screening. First, we screened all titles and determined for each paper if it was on the topic of selection. If we could decide from the title alone that this work did not cover the topic of a 3D interaction method, we excluded it from our list. In this phase, we removed 2,249 papers from our list, resulting in 567 entries. Second, we read the abstract of each paper. In contrast to the title-screening, where we would directly include or exclude works from our review, we used a three-point Likert scale (Agree - Neutral - Disagree). We re-evaluated all papers labeled with neutral in the first iteration of the abstract-screening, deciding for each neutral labeled paper to either be included or excluded. This resulted in 252 included papers after the screening phase from our query search.

In the last filtering phase of PRISMA, we read each remaining paper and tested for eligibility. We defined eight exclusion criteria Table 3.1, and removed all works which fulfilled at least one of them. In the following, we provide explanations for each exclusion criteria.

1. **Not in English:** The paper is not written in English, and thus, is not accessible.
2. **Survey paper:** We did not include survey papers in our dataset. Instead, if suitable and feasible, we would directly include the survey's references into our list. If our query search already found a reference, we did not include it again.
3. **Paper does not describe a selection method:** Similarly to our exclusion criteria in the title and abstract screening, if the paper was not on the topic of selection methods, we excluded it. This also includes works that were only focused on object manipulation techniques, not on the selection part of an object manipulation technique.
4. **Near field selection method:** We decided to focus this systematic literature review on distant selection methods. Thus, if the presented method was a near field selection method only, we excluded it.
5. **Hardware prototype description:** If the focus of the paper was a proposal for a new input device and only explained the technical properties of the device, we excluded it.
6. **Movement restriction:** If a selection method restricted the users' movement, e.g., through the usage of a regular mouse, we excluded the work.
7. **Preliminary work with full paper available:** For literature where our crawling found an initial non-archival version but additionally the resulting full paper, we excluded the initial work.
8. **Paper does not give enough information:** This exclusion criterion mainly filtered out literature being an initial non-archival version, as the authors did not provide enough information regarding the properties of their selection methods.

	Reason	Count
(1)	Not in English	4
(2)	Survey paper include references	9
(3)	Paper does not describe a selection method	38
(4)	Near field selection method	26
(5)	Object manipulation technique	15
(6)	Movement restriction	8
(7)	Preliminary work with full paper available	16
(8)	Paper does not give enough information	10

Table 3.1: Exclusion criteria for the eligibility phase of PRISMA with the amount of excluded paper for each reason.

3.3 Eligibility & Inclusion

From the found survey papers, we gathered 40 references but excluded 15 of them as our query search already found them. We also added the 5 works we found in our initial search on Google Scholar, which our query-search did not find. This resulted in 282 works. With our exclusion criteria, we excluded 136 works: 126 from our query results and 10 from the survey references.

This process left us with 126 papers from our database searching, 15 references from survey papers, and 5 works from our initial search on Google Scholar. In total, we included 146 papers for our systematic literature review.

3.4 Dataset Creation

For our systematic literature review, we decided to create a table with all our included works. In the process of analyzing the literature, we gradually extracted factors, with which we propose to classify and describe selection methods. One paper, most of the time, resulted in multiple different methods. This was either the case for literature which compared different selection methods, or if it compared one selection method in one of our extracted columns. This process resulted in 302 different selection methods. We only included methods that did not violate our previously stated exclusion criteria. Single methods were mostly excluded due to (4) *Near field selection method* and (6) *Movement restriction*. In the process of creating our dataset, we labeled each paper in 24 columns and extracted 10 different factors. We grouped these factors into fitting categories and determined suitable classes for each factor by investigating the literature.

4 Impact Factors for Designing Selection Techniques

This chapter reports and introduces the extracted design factors coming from 146 papers resulting in 302 methods. We propose four categories, in which we will describe each factor and provide our results. The categories are input and output devices, selection process, fine grain selection, and confirmation of the selection. Previous to these impact factors, we describe our findings regarding environmental factors, as different environments propose different requirements for selection methods. We provide tables with all citations of each category.

4.1 Environmental Factors

Different 3D environments affect the requirements of the selection method. Researchers tackled different challenges depending on the context in which the selection methods were proposed. In our investigation of the literature, we found that target density and occlusion impacted the capabilities of different selection methods. At the same time, the application design gave insights into the purpose of the proposed selection technique. For games and other immersive environments, where user experience and immersion are key factors, other challenges occur than for data analysis tasks. We found that study applications used either 2D or 3D environments independent of the used output device (see Figure 4.1 (a)). Furthermore, even in 3D environments 28.81% of methods only presented targets in one plane, thus, only studying a 2D selection space. 3D selection spaces are used when targets are placed in different depths of the environment. In the following section, we present our results regarding the application and setting of the literature.

4.1.1 Application

We labeled each method with the application or task used in the study to test the performance of the selection method. Our results can be seen in Figure 4.1 (c). We found that the majority of methods were evaluated with either a Fitts' Law task (111 methods) or by having the participants select arbitrary targets (107 methods). The class *immersive* contains 38 methods and includes applications such as selecting objects in a museum or game-like environments. Lastly, we also found 22 studies studying text input and 10 methods studying interaction with GUIs or menus. We understand that methods studying text input, GUIs, and the majority of Fitts' Law tasks can only have a 2D selection space, as the targets are arranged in one plane without disparities in depth. As there has yet to be agreed on a generalized 3D Fitts' Law task to easily compare the performance of 3D selection methods, it is currently difficult to compare the performance of selection methods. We found 13 methods with Fitts' Law tasks proposing a 3D selection task by placing their targets in various depths during one condition to create perspective distortion, which occurs in 3D environments [30, 88, 135,

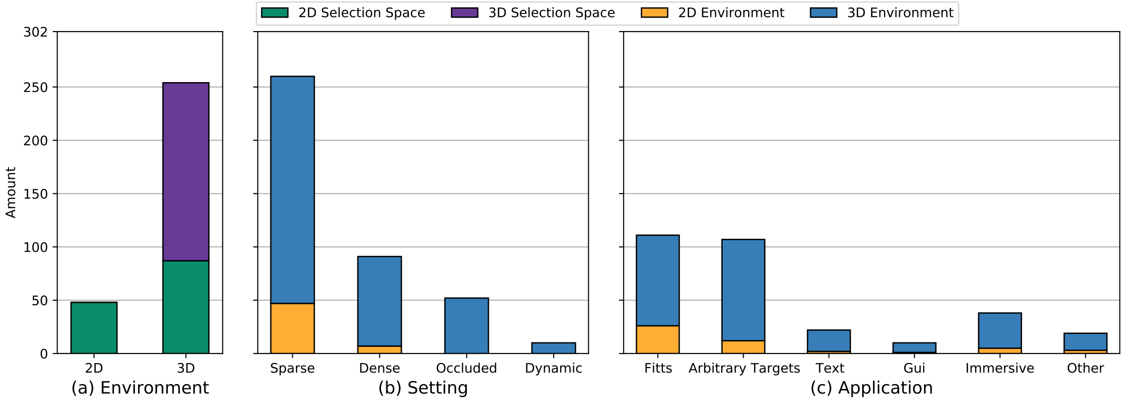


Figure 4.1: (a) Distribution of 2D and 3D environments and distinction between 2D and 3D selection space. (b) Amount of works testing different settings. One work can study multiple different settings. (c) Different applications, used in the studies. One study can use multiple applications.

169]. However, most researchers still used 2D or 1D tasks, as we found 88 and 10 respectively. To fully use the three dimensions of virtual environments, researchers used multiple arbitrary targets in different depths, which the participants then had to select [99, 195]. The benefit of this task is that the performance of the selection method can be evaluated for multiple different settings. Targets can be occluded and arranged densely. The disadvantage of these tasks is that comparing different studies is challenging, as target arrangement, size, placement, and density are not standardized but can highly affect the outcome. Methods studying text input and interaction with GUIs try to solve specific challenges. We found some literature proposing completely different methods of inputting text, which we could not classify as distant selection methods [33, 63, 157], but should be taken into account when building an application with a focus on text input. When using distant selection for text input, we found that the respectively best methods of six publications comparing different methods and reporting their words per minute (WPM) reached between 8.8 WPM and 16.43 WPM (M=13.15 WPM).

The application used in the studies is not an impact factor on selection methods per se. However, it gives insights into each method's goals and further clarifies why it is difficult to compare different methods with no standardized task. It also gives us a better understanding of which domains are interested in distant selection methods. We found that Fitts' Law tasks and arbitrary target selection are most widely used to test 3D input selection methods.

4.1.2 Setting

We classify settings into sparse, dense, occluded, and dynamic settings (see Figure 4.1 (b)). Sparse settings are the most widely used, as we found 260 methods using environments in their study where target density is low. We labeled Fitts' Law task to have sparse settings, as target density is not high, and each target can be easily distinguished. Furthermore, this is the usual setting for most immersive virtual environments. Thus, it makes sense that it is also the most widely investigated type of environment. The main challenges of sparse environments are usability, target selection speed, and low user fatigue. Dense environments are contrary to sparse environments, and we

Sparse	Dense	Occluded	#	Citations
✓	✗	✗	199	[5, 7–9, 11–15, 17, 19, 20, 28–30, 32–35, 37–40, 43, 45, 47, 49, 51–53, 56–59, 65–68, 71, 72, 74, 76, 78, 80, 82–84, 86, 88, 89, 91, 93, 94, 97, 99, 100, 102, 106, 108, 118, 119, 121, 123–127, 129, 131, 133, 135, 137–139, 141, 143, 144, 146, 148, 150–152, 155, 156, 162, 166–170, 172, 176–179, 181, 182, 184, 185, 192, 194, 196, 198, 199]
		✓	12	[4, 90, 122, 161, 188, 193]
	✓	✗	25	[10, 24–26, 46, 96, 101, 159, 191]
		✓	24	[44, 99, 142, 149, 158, 195]
✗	✓	✗	28	[2, 6, 27, 48, 63, 79, 85, 98, 136, 157, 160, 174, 186, 189, 190, 197]
		✓	14	[114, 116, 130, 175, 180, 187]

Table 4.1: Results of the setting-combinations used in the literature.

found 91 studies testing their selection method in a dense environment. If target density is high, single target selection becomes more complicated if the method's accuracy is not very high. This challenge becomes even more extensive for small and distant targets. High target density often leads to occlusion, although occlusion can also occur in sparse settings. Occlusion can only happen in 3D environments, as only then can targets be displayed in different depths. These settings bring even more challenges, as the interaction technique needs to use all three dimensions of the selection space to select targets behind other targets. Occlusion can be solved by either enabling the control of the z-axis or some other form of depth selection. We found 52 studies testing their method with occluded settings. However, none of the 3D Fitts' Law tasks included occlusion in their application, as the targets were just placed at different depths but without occluding other targets. As these different settings propose different challenges, researchers often singled out a specific setting and built their selection method specifically to solve these challenges.

Furthermore, we found four studies with ten methods studying dynamic settings [24, 25, 48, 185]. In game environments, dynamic targets are widespread. However, we did not find many studies using dynamic settings as an independent variable. The literature mostly agrees that selection time in dynamic settings is slower in contrast to static settings. Standard raycasting-based methods have a more significant drop-off in performance than techniques implementing disambiguation mechanisms. Especially behavioral disambiguation leads to similar performance and outperforms raycasting methods in dynamic settings [48].

To build a selection method, being able to solve challenges from multiple settings, some studies proposed to progressively increases the difficulty of selection by adding density and occlusion [99]. The different combinations of settings can be seen in Table 4.1. We found that 25.08% of the studies evaluated their selection methods in multiple different settings. Lu et al. [99] first compared a set of selection methods in a sparse environment and excluded methods, which had the worst performance. They then increasingly and systematically added occlusion and density to their studies. This refinement approach to single out the best method of a bunch was also done in some other studies [195]. Furthermore, frequently target density or occlusion was seen as independent variable [149]. Cashion et al. [24] built a framework to compare different selection techniques with target density as one comparative factor. They also propose that a perfect selection technique should solve problems from both environments.

In the following sections, we present categories and factors to classify selection methods. Different techniques have different advantages and disadvantages, specifically regarding selection performance in different settings.

4.2 Input and Output Devices

There is a wide range of different input and output devices. On the one hand, this gives consumers and developers many options to choose from, and variety can be great to find specific solutions for specific use cases. The technical side of VR technologies majorly improved in the last years. Current VR-HMDs mostly come with displays with high resolution and a high refresh rate. The Oculus Quest 2 has two 1832 × 1920 LCD panels, with a refresh rate of 120 Hz[1], while the Oculus Rift DK1 had a resolution of 640 × 800 per eye with 60 Hz[2]. Also, budget VR solutions, like Cardboard[3] profited from improving smartphone screens. Thus, immersing users in virtual environments is not hindered anymore from too low resolutions, and developers can entirely focus on solutions for challenges in how users can interact with the environments. The tracking accuracy of input devices also vastly advanced in recent years. On the other hand, this high amount of different devices can lead to confusion on both ends, which devices one should choose. Especially choosing the correct input device can be challenging. Distant selection methods do not have to be adapted to the input device anymore regarding its tracking accuracy. Designers have to think about what kind of interaction they want users to be able to perform and how many DoF an input device therefore needs. However, it is also essential for input devices to have high usability.

4.2.1 Output Device

Although this work mainly focuses on distant selection in VR, as described in Section 2.2 and Section 3.1, we also included other display types into our search query. We investigated how the used display type changed over the years. In Figure 4.2, we can see that since the release of modern VR-HMDs, stereo displays and CAVEs got mostly replaced in studies regarding distant selection. Additionally, we observed that the releases of consumer-ready VR-HMDs amplified the interest of researchers in distant selection methods, as new consumers found new challenges in different environments, and researchers had to find ways with the emerging technologies to satisfy their expectations. We found that 75.17% of our gathered selection methods were proposed since 2013. VR-HMDs were used to evaluate 138 methods, CAVE and stereo displays for 35 methods, AR for 66 methods, large displays for 37 methods, regular displays for 22 methods, volumetric displays for 4 methods.

The choice of which display to use primarily relies on the type of the application. VR enables true 3D environments, whereas traditional 2D screens can display an environment in 3D but without depth cues. In Table 4.2 the columns environment and selection space describe what kind of environment got displayed in each method and in how many dimensions selection was possible. VR environments immerse the user by replacing his vision with an entirely virtual environment. Thus,

---

[1] `https://xinreality.com/wiki/Oculus_Quest_2`

[2] `https://xinreality.com/wiki/Oculus_Rift_DK1`

[3] `https://arvr.google.com/cardboard/`

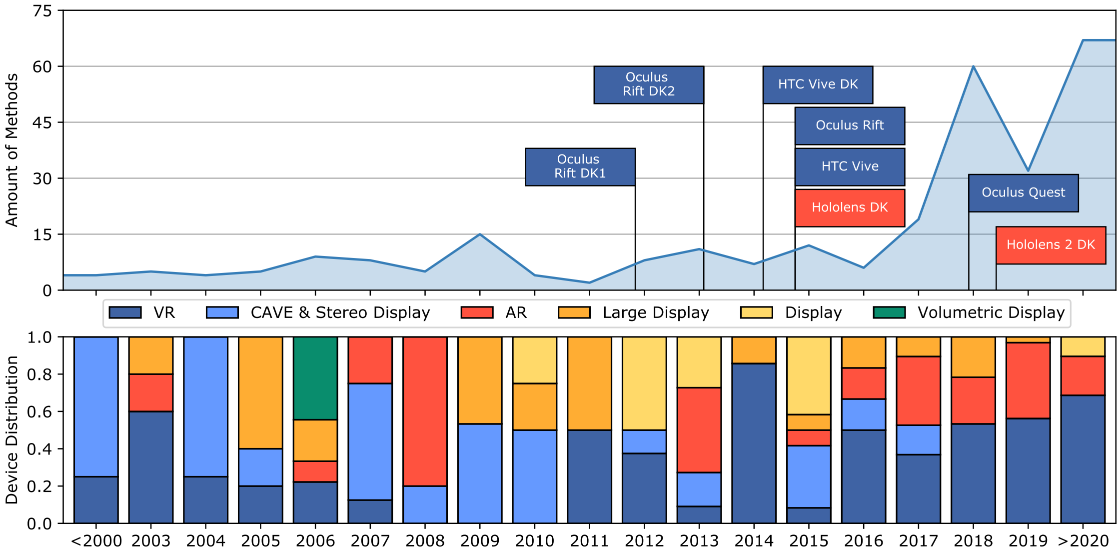


Figure 4.2: Amount of selection methods and distribution of output devices used in these methods over the last 20 years.

this virtual environment can only be 3D. However, the selection space can still be 2D. If the method did a study with a 2D-Fitts' Law task, the selection was restricted on one 2D plane, and thus, the selection space is 2D. The selection space can never have more dimensions than the environment space. 66.67% of methods using VR-HMDs utilize the 3D environment and provide a 3D selection space, meaning selectable targets can appear in different depths, which means that the selection method for such an environment needs to be able to select in 3D. Optical see-through AR can be very similar to VR regarding its properties of environment and selection space. Additionally, AR, especially video see-through methods, can provide users with 2D environments by projecting 2D planes into the environment for various use cases, e.g., interacting with text projected onto a wall similar [33]. Large displays also emerged, especially for providing a large space for information in collaborative tasks. We found two papers comparing distant selection in VR and AR. Both find AR does not lead to better performance in 3D selection tasks than VR, as one work found higher performance scores and higher levels of comfort in their virtual environment condition [126], and one found no significant effects between the environments [11]. While the most use of VR is for immersive entertainment content, the use case for large displays and AR are primarily informative. Figure 4.2 also shows that regular displays, as well as volumetric displays, mainly were used previously to the release of adequate VR and AR displays. We propose that interaction can converge to be very similar for all these display types and, thus, we treat the selection methods independently regarding its display type.

4.2.2 Input Device

We found eight main classes of different input devices. Different input devices differ in their amount of DoF, accuracy, immersion, and easiness of use. The most common input device for 3D interaction is a 6-DoF controller, where three DoF control translation and three DoF rotation. We understand that controllers can come in many form factors and shapes but achieve the same goal. Thus, we treat

Output Device	Environment	Selection Space	#	Citations
VR	3D	2D	46	[9, 12, 14, 15, 19, 27, 39, 44, 47, 59, 63, 76, 84, 89, 98, 126, 136, 137, 141, 151, 157, 166, 172]
		3D	92	[10, 11, 13, 30, 32, 43, 45, 56, 57, 68, 71, 88, 90, 99, 100, 108, 114, 118, 119, 122, 125, 127, 133, 135, 142–144, 149, 150, 152, 159–161, 167, 169, 180, 184, 188, 189, 191, 194, 195, 199]
CAVE	3D	2D	3	[198]
		3D	12	[2, 4, 6, 37, 48, 106, 158, 187]
Stereo Display	3D	2D	6	[5, 34, 65, 93, 94]
		3D	14	[8, 38, 82, 130, 156, 174, 175, 182, 186]
AR	2D	2D	5	[33, 139, 177, 192]
	3D	2D	32	[19, 28, 58, 83, 124, 126, 155, 185, 190]
		3D	29	[11, 17, 40, 52, 67, 72, 96, 116, 121, 123, 162, 170, 176, 179, 193, 196]
Large Display	2D	2D	32	[7, 20, 29, 49, 51, 53, 66, 78, 80, 86, 97, 101, 102, 131, 148, 168, 178, 197]
	3D	3D	5	[35, 74, 79, 129]
Display	2D	2D	11	[26, 138, 181]
	3D	3D	11	[24, 25, 85, 91, 146]
Volumetric	3D	3D	4	[46]

Table 4.2: Results for the different output devices, grouped by the dimensions of the environment and selection space.

all devices acting as a controller similarly. We identified 118 methods using a controller (including 14 wand methods) and 15 using a phone in the same way as a controller. Pham and Stuerzlinger [126] compared distant selection using a controller and a wand and identified the wand to achieve significantly better results regarding movement time, error rate, and throughput. Additionally, users found the wand more comfortable and easy to use due to its lesser weight and smaller shape. In their study, Pham and Stuerzlinger used an HTC Vive controller, which weighs 203g. However, more modern controllers significantly reduced the weight of their controllers (e.g., the oculus quest 2 controllers weigh 126g [4]), leading to more comfort of modern controllers. A strong benefit of controllers is the number of different modalities like buttons and joysticks, which wands miss. This is also a reason to use a smartphone as an input device. The touchscreen can dynamically give users the type of additional interaction possibilities they currently need. Another reason to use smartphones as input devices is that most users have their smartphones with them at all times. Thus, interacting with, e.g., large displays in public spaces can happen with one's own personal controller without the need for additional tracking systems [128]. Both inside-out and outside-in tracking often rely on camera systems to track the controller and gain their position. By placing dedicated markers on the controller, easy tracking can be allowed. Using the hands as an input device has many similarities with using a controller, as hands also have six DoF. When using hands as an input

---

[4] `https://www.vrfocus.com/2020/09/all-the-specifications-for-oculus-quest-2/`

Input Device	Input Device	#	Citations
Controller	None	118	[2, 4, 5, 9–15, 24, 25, 27, 30, 32, 34, 35, 37, 38, 43, 45, 46, 48, 49, 57, 63, 66, 76, 78–80, 86, 88–91, 96, 99, 114, 119, 126, 131, 137, 138, 141, 142, 149, 156–161, 166, 169, 172, 174, 175, 180, 182, 184, 186, 187, 189, 190, 194, 195, 198, 199]
Hand	None	38	[9, 17, 20, 26, 29, 40, 49, 51, 66–68, 82, 88, 93, 97, 101, 102, 108, 118, 121, 122, 133, 146, 152, 168, 178, 181, 188, 198]
Head	None	40	[19, 28, 39, 44, 57, 67, 72, 83, 98, 100, 123, 135, 137, 141, 143, 149, 157, 162, 184, 190–192]
	Controller	10	[35, 56, 58, 83, 96, 123, 137, 170]
	Hand	9	[4, 29, 65, 66, 74, 93, 100, 127, 196]
	Eye	4	[83, 150]
	Watch	3	[59, 71]
Eye	None	23	[8, 19, 26, 32, 53, 57, 83, 85, 119, 124, 135, 136, 144, 151, 167, 197]
	Controller	2	[83, 169]
	Hand	5	[6, 125, 181]
Phone	None	13	[28, 33, 52, 94, 130, 139, 148, 155, 176, 177]
	Controller	2	[114, 179]
Watch	None	7	[28, 129, 148]
Touchscreen	None	28	[7, 26, 28, 47, 52, 59, 84, 106, 116, 148, 155, 162, 177, 185, 193]

Table 4.3: Results of input device combinations used in the literature.

modality for 3D interaction with camera-based systems, additional markers are often needed to be placed on the user's hand. Other systems rely on hand pose detection via camera images, like the Leapmotion[5]. Another way of enabling hand tracking are glove-based systems, which rely on IMU data. We found 38 methods using hands as their primary input device. Four publications compared hand-based pointing with controller-based pointing [9, 49, 88, 198]. Across these studies, we found that hand-based selection methods never outperformed controller-based methods, and there is a trend for higher performance levels for the latter. Two studies found significantly higher performance for distant selection with a controller [9, 88], while the others found no significant effects on performance between these input modalities [49, 198].

We found 96 methods using head or eye gaze as an input modality. Gaze mostly only has rotational tracking, thus, resulting in three DoF. The positional three DoF can be added if movement of one's own avatar in the virtual environment is allowed. Eye gaze is very similar to head gaze in this regard, but the selection cursor is not bound to be in the center of the screen. Thus, different objects can be selected without moving the head, but eye gaze selection methods suffer from the Midas Touch problem [64], as mentioned in Section 2.1. Head gaze is always available when using an HMD to create a 3D interaction environment, while eye trackers are mostly not included and, thus, require additional hardware. Head and eye selection methods can be further sub-categorized into eye-root/head-root and eye-hand/head-hand techniques. The former being also called gaze

---

[5] `https://www.ultraleap.com/product/leap-motion-controller/`

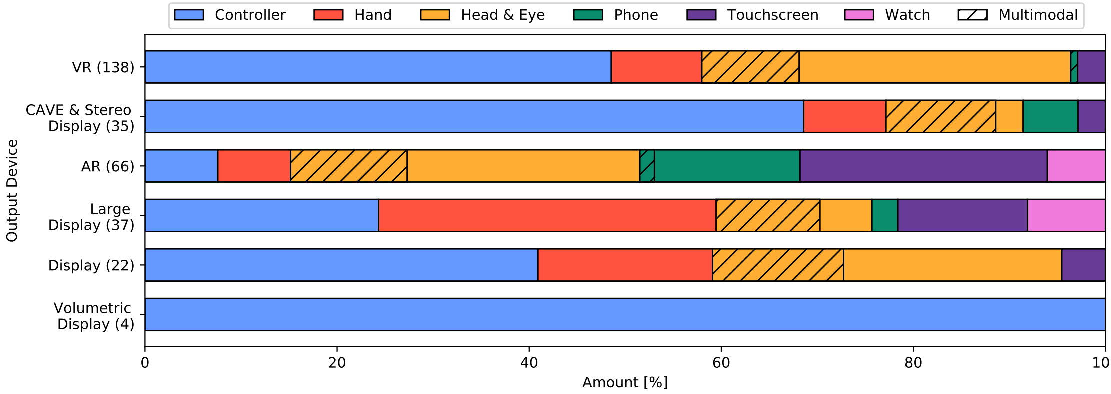


Figure 4.3: Distribution of input device usage for each output device in the literature.

techniques. We found 40 and 23 head and eye gaze techniques, respectively, and 26 and 7 head-hand and eye-hand techniques. Across the literature, we found five works comparing head-gaze and eye-gaze selection performance [19, 57, 83, 135, 150]. The studies do not conclude the same results regarding the performance or user preference. One work did not find any significant effects on task completion time or accuracy [150], two works found faster completion times for head gaze [57, 135], and two for eye gaze [19, 83]. Accuracy is also not conclusive, as the works which found faster completion time for eye gaze did at the same time achieve lower accuracy. Two works found lower error rates for head gaze and one for eye gaze. User preference was also different across the literature. We found three works comparing gaze techniques with eye-hand/head-hand techniques [100, 123, 137], where the general consent is that the latter methods outperform gaze-based methods. We lastly want to provide results regarding the performance of controller as input modality versus head and eye input methods and found 11 works in the literature doing that comparison [28, 32, 119, 123, 137, 141, 149, 157, 170, 184, 190]. We found that using a controller as input modality significantly outperformed head-based methods in eight studies, was worse in two studies, and equal in one study regarding selection time. The error rate was significantly lower for a controller in four studies. Four studies did not report significant effects. In five studies, subjective feedback preferred the controller, while five did not report significant effects on user preference, and one did not report subjective feedback results. From this, we propose a clear trend towards better performance and usability for controller-based input methods.

Another thing to consider when building a 3D interaction technique is bimanual input. This is possible when using hands or 6-DoF controllers as an input device. When using bimanual input, each hand can either select targets independent of each other [157, 160] or both hands perform one single selection [122, 188]. Designers need to take into account that performance with the non-dominant hand might be different than with the dominant hand [35]. However, handedness does not impact pointing-based distant selection methods, as the performance was very similar, independent of the hand. In contrast, handedness did impact stylus or mouse selection performance [49]. Multimodal input also got used for different distant selection methods (see Figure 4.3). On the one hand, head-finger techniques are multimodal, as two input devices (head-tracking and hand-tracking) are needed to gather two positions. We also found methods combining a phone/tablet with a wand to increase the DoF for selection.

As seen in Figure 4.3, for each output device, a different distribution of input devices are mainly used. While in virtual environments, the majority (48.55% for VR-HMDs, 68.57% for CAVE & Stereo Displays) of methods use a controller as an input device. In AR, most methods use head/eye-pointing. However, most current smartphones can also create AR through video see-through. Hence, the phone and touchscreen are also often used as input devices. We differentiate between phone and touchscreen as input when selections are performed by moving the phone or targets on the display are selected via touch. In contrast to VR, AR does not put the user in a completely different environment but rather augment reality. A hand-held controller disrupts regular interaction, and, thus, hands-free approaches [192] are more usable for AR. Interaction with large displays, in the context of its use-cases and, thus, usable input devices, can be compared with AR. Literature researching smartwatch-based input [28, 129, 148] show that interaction with these 3D environments can be performed with daily worn devices, thus, achieving the goal of being able to perform an interaction with these output devices at all times. With virtual environments, ones' sensory experience gets fully replace by a digitally created one, intending to immerse them into the virtual environment fully, such that they have a sense of presence in *being there* [54]. Interaction is only happening with this digitally created environment. Thus, having additional controllers to interact with the environment does not negatively affect interaction options, as there are none. Controller, although becoming lighter, can still affect user fatigue more than just using the hands. By enabling more DoF in interaction with hand interaction only, through multiple natural gestures, hand-based input methods could replace controller interaction at some point.

To generalize our findings regarding used input devices, we propose that currently, there is a strong correlation between output devices and input devices. Through a better understanding, which 3D selection methods work best in different environments, input and output devices might become an independent factor.

4.3 Selection Process

This category includes the factors, which classify the process of performing the selection. A user needs to control a cursor with an input device towards a selectable object for a selection to happen. The positioning, visualization, and input & mapping explain how the cursor is controlled, designed, and mapped to the visual space. Again designers have to weigh between usability, selection time, and accuracy. Different settings affect how difficult performing a selection is and, thus, how much support a user needs to perform it.

4.3.1 Positioning

The factor positioning describes the relationship between movement in the physical space and the digital space. If the virtual selection technique directly connects to a user's movement in the real environment, positioning is absolute. Absolute positioning mimics interaction with the real world and, thus, often uses the humans' natural ability to point at objects. In contrast, relative positioning uses mapping from movement in the physical space to translate it onto moving the cursor in the digital space. The cursor can still have a direct connection to the input device by being tracked at

XY	Z	#	Citations
Absolute	None	219	[2, 4–6, 8, 9, 11–13, 15, 19, 24–30, 32–35, 38, 39, 44, 46–48, 52, 53, 56–59, 63, 65–68, 71, 72, 74, 76, 78–80, 83–85, 89, 90, 93, 94, 96, 98–102, 106, 108, 114, 116, 118, 119, 121, 123–127, 131, 135–139, 141, 143, 144, 146, 148–152, 155–159, 162, 166–170, 172, 176–181, 184–199]
	Absolute	26	[13, 14, 17, 45, 46, 82, 88, 114, 130, 142, 158, 160, 161, 169, 174, 175, 182]
	Relative	17	[10, 37, 40, 43, 67, 88, 99, 108, 122, 129, 133, 156, 193, 195]
Relative	None	40	[7, 20, 26, 28, 49, 51, 59, 76, 86, 91, 93, 97, 123, 137, 148, 157, 162, 169, 177, 178, 185, 199]

Table 4.4: Results of the positioning (XY) and depth positioning techniques (Z) in the literature.

all times, but can also be rate-controlled with, e.g., a joystick [164]. We can categorize these two different relative positioning methods into direct and indirect techniques [183]. Table 4.4 depicts our results.

Positioning in 3D environments can be split into two different parts, movement of the cursor left, right, up, and down (XY) and depth movement (Z). While positioning is the only form of movement control in 2D environments and is necessary for interaction with any environment, depth-positioning is not directly necessary in sparse settings. Most selection methods simply select foremost target in 3D environments [72, 90], or have the cursor at a fixed distance [13, 14, 88]. Both positioning and depth-positioning can use absolute or relative positioning independent of each other.

We found 262 methods with absolute positioning XY-direction. Except for some refinement techniques, absolute positioning methods almost always have a 1-to-1 control-display (CD) ratio. For absolute methods, we differentiate between forward-ray methods and two-point-ray methods (see Section 4.3.2). However, both methods have a direct relationship between the users' motor space and visual space. Touchscreen-based input is also mainly absolute [164] if it is not used like a trackpad [26]. In absolute positioning methods, the cursor's movement is continuously dependent and directly coupled to the movement of the input device in the real world. As the physical boundaries limit humans' direct control space, distant selection methods extend this space by taking the pointing direction. The pointing metaphor is the closest technique mimicking real-world interaction in virtual environments. However, also controller pointing methods, acting similar to a laserpointer [66], are classified as absolute, as they directly transfer the users' movement to the output device. Difficulties of absolute positioning methods are jitter [120] and inaccuracy of humans at distant pointing [6, 104, 105]. To counter jitter, filtering methods can be applied. In the literature, we found methods using the Kalman filter [118], 1€ filter [10], and lowpass filter [178]. Baloup et al. [10] report that when filtering the ray, participants made 50% fewer errors.

We found 39 methods with relative positioning in XY-direction. These methods control the cursor's position by direct mapping or indirectly moving the cursor with input. The CD-ratio for relative positioning methods does not have to be 1-to-1 or even linear. Frequently it is helpful to have a dynamic CD-ratio to mitigate the smaller control space and still be able to interact with the whole environment [178]. As a virtual environment usually span out over 360 degrees, and users can quickly look around by turning their heads, in theory, it is possible to point at objects without looking at them. When using a relative positioning technique, the control space of the controller needs to be very large, or users need to recenter the device often, which is how we usually interact with touchpads. In relative positioning methods with hands as the input device, often clutching is used for this [51, 178]. Also, when using a mouse, users can only interact and look in the general direction, where the mouse is located in the real world, which restricts the user from interacting with 360 degrees of the environment, so we excluded these methods. In contrast to relative positioning in 2D environments, here, the input device either controls three DoFs [33], or a 2D position is mapped onto the virtual space, and a ray casts from the 2D position orthogonal to the viewport [126]. Furthermore, we found methods that mapped a 3D space onto a 2D selection plane [137]. The cursor on the image plane is then, in this case, controlled relatively. Movement in the real world does not directly match with the position of the virtual selection tool. For immersive virtual environments, developers want to create the highest possible level of immersion and presence. Frequently in VR, users have a virtual representation of themselves, a virtual avatar. To achieve high levels of immersion and presence, the position and movement of the virtual avatar should match with the movement of the user in the real world to enable self-location [16, 73]. It has even been shown that the appearance of the avatar can affect pointing performance [145]. For absolute distant pointing, the direction from the users' body gets extended, usually via raycasting.

In the literature, 12 researchers compared relative positioning and absolute positioning methods [26, 28, 59, 76, 123, 137, 148, 157, 177, 178, 185, 199]. The results show that absolute positioning methods led to better performance than relative positioning in six studies. There were no significant effects in four studies, and two studies found relative positioning to perform better than absolute. Both studies with better relative positioning performance compared smartwatch-based input devices. Relative positioning methods, however, often led to lower error rates.

While absolute positioning seems to be favored to control the XY-direction, it is not easily usable to control the Z-direction. We found 43 methods, allowing positioning in the Z-direction, where 17 are relative, and 26 are absolute. Positioning in Z-direction is mainly needed to deal with occluded targets by controlling the depth of the cursor. As distant targets can be arbitrary far away, many targets would be out of reach. Furthermore, humans are least precise in controlling the depth of the cursor [92]. The GoGo technique was a pioneer in the development of distant selection techniques, as it allowed the selection of distant targets by extending the virtual arm relatively by moving ones' own hand forwards [133], in a time where interaction with virtual environments was still mainly done with virtual hand techniques for near target selection only. Due to a CD-ratio larger than 1-to-1, targets out of reach could now be selected. Many following researchers built upon this technique or used it for comparison with their selection methods while using both linear and nonlinear CD-ratios [88, 95, 156, 193]. Other authors proposed ideas on how to use absolute positioning for the Z-direction and still be able to reach distant targets, e.g., using a fixed offset [88].

Controlling the cursor with absolute positioning might lead to a more immersive experience, as self-location and, thus, embodiment can reach higher levels. The user does not control the cursor with a disconnected tool but can directly interact with these objects by extending the selection tool for distant targets. However, relative methods have often been shown to be more precise than absolute pointing methods. If a distant target selection method wants to include depth-positioning, either relative positioning is needed with a suitable CD-ratio, or a fixed offset is needed. The latter has the disadvantage that even though targets in the distance can be selected, the user is limited in selecting targets only at that specific distance.

Input	Origin	#	Citations
Forward Ray	Controller	118	[2, 4, 5, 9–13, 15, 24, 25, 27, 28, 30, 32–35, 37, 38, 43, 45, 46, 48, 49, 52, 56, 57, 63, 66, 76, 78–80, 88–90, 94, 96, 99, 114, 119, 126, 131, 137–139, 141, 149, 155–161, 166, 169, 172, 174–177, 180, 182, 184, 186, 187, 189, 190, 194, 195, 198, 199]
	Head	71	[8, 19, 26, 28, 32, 39, 44, 53, 57, 67, 72, 83, 85, 98, 100, 119, 123, 124, 135–137, 141, 143, 144, 149–151, 157, 162, 167, 169, 184, 190–192, 197]
	Hand	28	[9, 17, 29, 40, 49, 67, 68, 82, 88, 101, 108, 118, 121, 125, 129, 133, 146, 152, 168, 178, 181, 188, 198]
	Touch	17	[47, 52, 84, 106, 116, 148, 155, 185, 193]
Two Point Ray	Head/Body	22	[4, 6, 29, 35, 58, 59, 65, 66, 71, 74, 93, 96, 100, 123, 127, 170, 181, 196]
	Hand(held)	4	[102, 122, 179]
Origin Offset	Hand(held)	6	[13, 14, 88, 142, 169]
Direct Mapping		23	[7, 20, 26, 28, 51, 76, 86, 91, 93, 97, 130, 148, 157, 178, 185, 199]
Indirect Mapping		13	[20, 26, 28, 59, 137, 148, 162, 177, 199]

Table 4.5: Results for the literature categorized by the input and grouped by the origin.

4.3.2 Input and Mapping

In this section, we look into how different selection techniques calculated the position of the virtual selection tool. As mentioned, absolute positioning methods are mostly raycasting-based. We can divide them into forward-ray methods, which use one point and a direction, and two-point-ray methods, which calculate the direction between two points. Head-Hand raycasting-based methods are the most known two-point-ray methods. Furthermore, we found origin offset methods, where positioning is absolute but with an offset in any direction, to select distant targets. Relative positioning methods mainly calculate the cursor position by mapping a position from an input device in the control space onto the digital space with a specific CD-ratio. Table 4.5 present our results for the input and mapping factor.

The large majority of studies we found when analyzing the literature used raycasting-based methods. Cone-based techniques are very similar to classical raycasting but often need some form of disambiguation mechanism. We found that 234 methods use forward ray casting and 26 two-point-ray methods. Raycasting methods offer an intuitive way to interact with 3D environments, as they reduce the problem of controlling all dimensions of the cursor and instead just cast an infinite ray, selecting the first target seen. Forward-ray methods are either cast from the head or an external input device, which can be a handheld or the hand itself. We also classify touch-based AR video-see-through methods as forward-ray methods, as a forward vector gets cast from the touch towards the target. The former has the advantage of being a hands-free method, which is desirable in some areas of 3D interaction. AR, for example, widely uses head-forward raycasting or budget VR devices like Cardboard. However, turning the head for every selection can be fatiguing, as well as the need to always look in the direction where users want to select targets can be cumbersome. Literature comparing head-forward and hand(held)-forward raycasting methods agree that hand(held)-raycasting outperforms head-forward raycasting significantly in almost every case [32, 67, 119, 137, 141, 149, 157, 190]. Also, subjective feedback concerning fatigue and easiness is better, as hand-based methods might lead to a better spatial relationship and, thus, seem more concrete to the user [157]. Some authors, however, state that device-free methods might be more practical in specific scenarios like outdoor AR, for example. Forward-ray methods are also relatively simple to develop and use, as only one input device is needed to perform raycasting. A controller can be held at wrist-level and, thus, lead to very little fatigue in contrast to head-hand raycasting.

Instead, two different points, gathered from two input devices, can calculate the direction. We found 26 two-point raycasting methods. One common idea is to use eye-hand raycasting, which mimics human pointing in the real world [6, 71, 93]. The direction can also originate from the fingertip instead of the hand [29, 93, 100, 181] or the arm with a smartwatch [59]. The origin can also be fixed at the chin [58, 123], the belly [66][6] or at an arbitrarily fixed point [69]. However, also non-head-based methods experimented with raycasting methods from two different positions. Matulic and Vogel [102] explore different hand-rooted methods by casting rays between different joints of the finger. They found that rays calculated between the most distant finger-joints are more stable and perform better than rays calculated from joint combinations with less distance. However, we found no comparisons between finger-joint methods and classical hand-based methods. Literature comparing forward-ray and two-point ray methods mostly agree that forward raycasting outperforms two-point raycastin [4, 29, 66, 96, 100, 123], as three methods performed better, two equal, and one worse. However, note that the compared techniques here were party vastly different, and not many studies compared similar two-point and forward raycasting methods. Head-hand-based methods have the advantage of mimicking pointing in the real world and not leading to the eye-hand visibility mismatch problem [4], but also have the same disadvantages of not being able to select occluded objects. Raycasting techniques, where both the origin and direction of the ray originate from one input device, are less fatiguing to use, as users are not forced to hold their arm in a pointing position but can cast the ray from anywhere [92].

Origin offset methods enable distant target selection at one specific distance. These mostly use absolute positioning and absolute depth positioning and can either mimic classical virtual hand techniques or use a fixed-length ray [13, 88]. We only found six origin offset methods. However, we can state that these techniques are not applicable on every distant target selection task due to their restriction to only select targets at one specific location.

Mapping-based cursor positioning can be divided into direct and indirect mapping [183]. Direct mapping approaches are similar to absolute positioning, as they also have a constant connection between motor space and visual space, however not with a 1-to-1 CD-ratio. We found 23 direct mapping-based techniques. These range from methods like distant free pointing, where the hand controls the cursor [178], over VR mouse emulation [76, 199], to trackpad based approaches [7, 185]. Many of these direct mapping-based approaches were developed for large displays. These relative positioning methods very often have higher accuracy and precision, in contrast to raycasting [178]. Indirect mapping approaches do not have a constant connection between the input device and the cursor. We found 13 methods using indirect mapping. This mapping approach is used in trackpad techniques [26, 28, 148] and joystick controll [137, 199]. Boonnak et al. [20] implemented both a direct and indirect mapping approach based on the hand-gesture approach. The advantage of indirect mapping approaches in contrast to direct mapping is that they can move the cursor in an arbitrary large environment. They only move the cursor in a direction and do not have a direct

---

[6] Note: While the authors state that belly / index-finger pointing is a fixed origin pointing technique, this is only true due to their study setup in which the participant was not allowed to move. In a 3D selection context, this can be considered as normal belly / index-finger pointing.

Category	#	Citations
Cursor	194	[5, 7, 8, 11, 19, 20, 24–26, 28–30, 32, 33, 35, 37, 39, 44, 49, 51, 57, 59, 63, 65–67, 71, 72, 74, 76, 78–80, 82–84, 86, 88, 91, 93, 94, 96–102, 114, 118, 119, 123, 124, 126, 127, 129–131, 135–139, 141, 143, 144, 146, 148–151, 157, 160–162, 168–170, 174, 175, 177–180, 182, 184–187, 190–197, 199]
Ray	57	[4, 6, 10, 12–15, 27, 32, 34, 35, 38, 43, 46, 52, 56, 58, 68, 76, 85, 89, 90, 96, 99, 119, 123, 125, 141, 156, 158, 166, 172, 174–176, 180, 188, 195, 198, 199]
Spline	7	[2, 48, 99, 122, 159, 189]
Cone	13	[9, 24, 99, 108, 121, 158, 193, 195]
Virtual Hand	17	[17, 40, 45, 67, 99, 108, 133, 142, 152, 156, 181, 193]
None	14	[47, 52, 53, 83, 106, 116, 155, 167]

Table 4.6: Results for the visualization of the selection tool.

relationship. However, depending on the size of the environment, this can take a long time, or when higher CD-ratios are used, be inaccurate at small target selections. We did not find conclusive results when comparing direct and indirect mapping approaches regarding performance.

4.3.3 Visualization

The visualization of the selection technique has a fundamental impact on how the selection method works. The chosen output device and dimensions of the environment have a substantial impact on the visualization, as the cursor also has to have three dimensions in a 3D environment, in contrast to 2D environments [5]. In 3D environments, it can help users gauge the depth of the cursor. The visualization can be decoupled from the input method, meaning a raycasting-based method does not necessarily use a ray visualization. Most methods we found used a cursor as a visualization technique. However, we found multiple different shapes of cursors. Table 4.6 shows our results of the literature using different visualization techniques.

We found 194 methods using a cursor visualization. This includes crosshair visualizations [169, 175, 177], volume cursors [175, 186], circles [83], and others [187]. This cursor usually is rendered at the target closest to the user in the pointing direction. With depth positioning, the user can control the depth of the cursor. The literature mostly decided to use non-directional cursors, e.g., a circle or sphere, in 3D environments, as this does indicate that the cursor does not have a preferred orientation [161]. 3D cursors should be preferred over 2D cursors, as they help users to approximate the distance from the cursors towards targets [5]. Volume cursors sometimes use the cursor to enable the selection of occluded targets, as the cursor makes occluding targets transparent [174, 175, 186]. Vanacken et al. [175] state that both a 3D volume cursor and depth ray outperform a simple 3D point cursor.

Some researchers provide a visualization of the ray to helps users get a sense of the pointing direction. Here, infinite ray representations achieve better performance than finite rays [13], as with the finite ray, users need to use depth cues to move the cursors onto a target, adding one more level of complexity. We found 57 methods using a ray visualization, where 96.45% come from either forward-ray or two-point-ray input. Other methods like depth ray include a movable controller on the ray, as they include depth positioning [46, 174]. Additionally, this ray can also be curved and visualized as a spline or cone. Spline visualizations almost exclusively use some form of heuristic to select targets close to the ray, and stick the end of the ray to the target. In terms of technique, spline visualizations are similar to cone techniques but with a different visualization. Spline visualizations clarify that only one target can be selected, and cone visualizations often deliberately select a subset of targets. However, cone techniques can also work as a single-target selection technique and help the user select small and distant targets with their large volume in contrast to a ray.

Only a few distant target selection methods rely on using only a virtual hand as a visualization technique. Virtual hand techniques like GoGo move the virtual hand into the depth of the environment [133] or even extend the arm visually [40]. Ninja Hands shows multiple virtual hands in the environment [142]. However, using only the virtual hand as a visualization tool is often insufficient to select distant targets. Video-see-through AR uses a smartphone or similar to create a 3D environment. As interaction here is happening directly with the touchscreen, oftentimes, no visualization is given.

Although the visualization of the selection tool does not directly affect the selection process, it can support users in their selection by giving additional depth cues. We did not find studies comparing one selection technique with multiple different visualizations except for a study by Argelaguet and Andujar [5]. They state that although a visualized ray should help users by showing the direction they are pointing at, ray visualizations are less accurate than cursor representations. However, a visualization should also match with the used technique, and support the users to understand what options they have when interacting with this technique.

4.4 Fine Grained Selection

Target selection in dense and occluded environments is a lot more challenging than in sparse environments. Selecting a specific target in dense settings is often challenging, especially if targets are small or very far away. In occluded settings, some targets would no even be selectable with standard raycasting methods, as these always select the first target hit by the ray. One technique, we previously mentioned in Section 4.3.1, to deal with occluded targets is to enable controlling the cursor in Z-direction to gives users manual control over the selection. Other techniques help the users by extending the size of the selection tool like cone or volume cursor techniques. However, with the latter, often, multiple targets would be selected by the selection tool. Therefore, a disambiguation mechanism can decide which target should be selected. Refinement strategies can also deal with dense and occluded environments by giving users more fine granular control over the cursor when close to a target or introducing a second stage into the selection. This section presents our findings on different methods to deal with challenges given by dense and occluded environments, categorized with the factors disambiguation & refinement, multi-target selection, and depth selection.

4.4.1 Disambiguation & Refinement

Ambiguity can either be caused due to the environments being dense and occluded or because of the used selection technique [147]. Some raycasting-based methods classify themselves as cone techniques to select multiple targets at once. Mechanisms to help with the choice of which target should be selected are widely called disambiguation mechanisms [3]. The usage of additional

Disambiguation	Stages	#	Citations
Manual	≥ 2	54	[2, 17, 24–28, 45, 46, 78, 79, 83, 85, 90, 99, 106, 108, 114, 116, 144, 148, 160–162, 180, 193, 195, 197]
Heuristic	1	22	[8, 24, 37, 99, 101, 108, 121, 123, 125, 142, 149, 158–160, 174, 175, 186, 193]
	≥ 2	7	[10, 24, 25, 158, 187, 189, 193]
Behavioral	1	5	[46, 48, 56, 143, 152]

Table 4.7: Results for different disambiguation methods, grouped by the amount of steps the method needs to select a target.

progressive refinement of the selection can support users to perform an accurate selection. We understand that refinement techniques can always be labeled as two-stage processes. Heuristic approaches are mainly used with cone or volume cursor and automatically select targets depending on a heuristic, e.g., their distance to the cursor. Lastly, also intent prediction can be used to detect which target the user wants to select. A default selection without disambiguation or refinement only selects a target that is directly hit by the ray. We found 88 selection methods using either disambiguation and refinement processes (see Table 4.7).

Using a cone as a selection tool can have multiple advantages. Due to the larger selection tool, the cone can counter hand jitter and inaccuracies and, thus, make it easier for users to select small or far away targets. With a cone, the selection cursor gets bigger the further away it is getting cast. However, with a cone, frequently, multiple targets get selected. Disambiguation mechanisms like a heuristic select the closest target to the center of the cone, and refinement techniques use a second stage to clarify the selection. By implementing the heuristic raycasting-based methods as ray splines, the ray is "magnetically drawn" towards targets with a threshold [159]. Two hands can also control ray splines, to control the bending of the ray, to selected occluded targets [122]. The majority of the heuristic methods select the target with the smallest distance to the cursor. The most widely used distance metric is calculating the euclidean distance, but also the angular distance is used and found to outperform Euclidean distance in some cases [99, 187].

Heuristically one-stage disambiguation mechanisms, e.g., the bubble cursor [174, 175, 186] widely uses sphere cursor visualizations. However, also ray techniques can be adapted with a heuristic and then be drawn as a ray spline [99, 159]. Heuristic methods outperform their manual counterparts, especially in occluded and dense settings [99, 174]. Instead of a heuristic, using intent recognition and prediction of targets can also help users at target selections [56], which can help, especially to select small targets [48]. Due to the limited number of studies using behavioral approaches, and their ambiguous results and lack of comparison with classical raycasting, we can not make a definitive statement about their usefulness for distant target selection, especially as they, even more than heuristic disambiguation, have the disadvantage being a black-box to the user.

A widely used method to deal with dense settings are refinement strategies. Frequently refinement strategies are used, either by giving the user more fine granular control over the cursor, when close to a target [26, 83], or by first choosing a subset of targets to then manually select the desired one [25, 195]. While none of the methods combining absolute and relative positioning significantly outperformed classical raycasting regarding selection time, the authors found that error rates went down significantly [10, 26, 83]. Manual refinement strategies often display all selected targets in a second stage. Grid wall [195] places all targets hit by a cone in a grid [195]. SQUAD [25, 79]

Disambiguation	Depth-Positioning	#	Citations
✓	✓	14	[10, 17, 45, 46, 114, 142, 158, 160, 161, 174, 175]
	✗	31	[24, 25, 46, 79, 85, 90, 99, 108, 114, 116, 149, 162, 180, 187, 193, 195, 197]
✗	✓	19	[13, 14, 40, 43, 46, 67, 99, 108, 122, 129, 130, 133, 156, 174, 175, 182, 193, 195]
	✗	6	[5, 85, 180, 188, 195]

Table 4.8: Results of methods, which implemented depth selection, either through depth-positioning, disambiguation, or other methods.

uses cone selection first to select a subset of targets, to then place it in four parts of the screen for a more accessible second refined selection. Refinement approaches help in dense settings, as users struggle with accurate selections on small and distant targets. The usage of two-stage approaches can lower error rates and circumvent miss-selections. Refinement strategies usually outperform classical raycasting in dense settings while being slower in sparse settings [25, 108, 116, 193]. Starfish cursor [187] is a heuristic disambiguation technique that targets close to one pre-selected target can all be selected based on different heuristics. This is visually shown by having the cursor look like a starfish, where the legs spread out into different directions.

Disambiguation methods with closest distance heuristics can be great for small distant targets. However, in very dense settings, they can lead to confusion and user frustration if multiple targets are in the same proximity and the user does not have complete control over the target selection. It is also crucial that users know and understand which target they will select when pressing the trigger. Spline visualizations and cone visualizations often use a closest heuristic; however, splines visualizations select only one target, while cone visualizations often lead to a refinement stage. Refinement strategies are great to deal with these high-density settings by still letting the user exactly chose which target should be selected, but with refinement steps. Disambiguation mechanisms can help speed up selections for sparse settings, as the selection tool does not have to line up with the target perfectly. Developers have to consider their setting to decide if the extra time needed for a refinement compensates for the additional necessary step in contrast to a manual or heuristically one-stage method. However, by automatically switching between a refinement strategy and heuristic disambiguation depending on the setting, one can use the advantages of both methods [24].

Many refinement strategies work by first selecting a preset of targets and then refining the selection by choosing one of the displayed targets. Furthermore, we also found methods explicitly to select multiple targets. The application of these methods is primarily environments with data points. Stenholt [160] transferred well-known multi-target selection methods from 2D applications into 3D interaction (brushing, lasso, and magic wand). Montano-Murillo et al. [114] used slicing volume techniques to select whole slices of a 3D environment.

4.4.2 Depth Selection

To fully utilize interaction in 3D environments enabling depth selection can be immensely beneficial. Depth selection, in this case, describes methods, which explicitly allow the selection of occluded targets, as most raycasting-based selection methods already allow target selection in different depths. Dealing with occlusion can either be achieved by enabling positioning of the cursor in the

Z-dimension. Furthermore, disambiguation mechanisms can also help in the process of selecting targets otherwise occluded. During our analysis of the literature, we found additional methods dealing with occluded settings. In Table 4.8 all methods, which implemented depth selection can be found.

In virtual environments, movement of one's own avatar and, thus, changing the viewport is often allowed [85]. By re-positioning the camera targets, previously occluded might now be selectable. The World in Miniature approach allows users to re-position their avatar anywhere in the environment with the help of a miniature version of the environment in their hand [161]. However, in some applications, it might not be desired to move around. The majority of the literature implemented depth selection with disambiguation methods, mainly two-step refinement mechanisms. Depth Ray combines refinement with depth positioning by locking a ray in a first step and then allowing the control of a cursor on this ray [46, 116, 128, 175]. These methods have been shown to outperform non-heuristic approaches in occluded settings [175]. Bimanual interaction techniques can also allow depth selection by performing a selection where the rays cast from each hand cross [188]. Manipulating the environments to uncover occluded targets can also be used [5, 195], but it might not be usable in immersive games.

In occluded settings, refinement strategies allow selecting targets, impossible to select with classical raycasting without depth positioning. One example is to give users a mirror to place in the virtual environment to selected otherwise occluded targets [85, 90]. Another way of dealing with dense or occluded environments is to manipulate the environment so that the setting becomes sparse locally. Either zooming towards a selected location can do this [2, 5, 17], or by pushing occluding targets away to unclutter that area and reveal occluded targets [116, 180, 195]. Another idea is to use slicing volume techniques to select one 2D slice of a 3D environment [114] or similarly tiltcasting [130] where slices of the environment are displayed on a smartphone, which can be moved through the 3D environment. Environment manipulation techniques can uncover occluded targets, e.g., Drill Pointing [116] and Smash Probe [195]. Shadow cone is a continuous refinement approach, where all objects inside a cone are selected and deselected as long as a user is holding down a button [158].

However, depth selection can also be enabled directly through depth-positioning. One example is the virtual hand technique GoGo [133], which has been proposed in 1996 and used in many following studies [88, 95, 156, 193]. The GoGo immersive interaction technique lets users control their virtual arms' growth relatively, thus, enabling the selection of distant targets. With the flexible pointer [122], both hands control a spline through the rotation of both hands, such that the user can navigate around targets to select occluded ones. Ninja Hands [142] is a virtual hand technique based upon ninja cursor [77], where the real hand controls multiple virtual hands in different positions at the same time. Origin offset pointing techniques are also able to select occluded targets [13, 14, 88]. Raycursor [10] is similar to depth cursor; however, the cursor can automatically detect occlusion. Let's go there [152] combines voice with pointing and allows semantic target selection also for depth, by being able to say "chose the furthest target".

Similar to the thought process, if using a disambiguation mechanism speeds up selection in dense environments, developers have to think about this when building a selection technique for occluded targets. For depth selection, two general options arise, with disambiguation & refinement strategies and depth positioning. While manual depth positioning gives the user direct control, users struggle most with gaging and controlling the depth of a cursor [92], and refinement strategies often require two stages, slowing down selection.

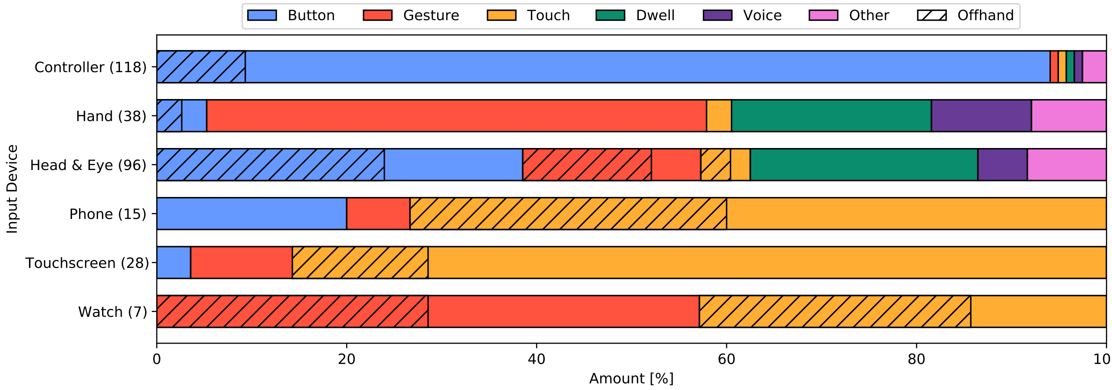


Figure 4.4: Distribution of used triggers for each input device.

4.5 Confirmation of the Selection

In the previous sections, we proposed factors to classify the indication of objects. Thus, this final category of our results deals with the confirmation of selection. Depending on the input device used, different possibilities arise. With the factor trigger, we describe the different modalities used to confirm a selection. The visual representation of the selection tool affects the users' perception of which target is going to be selected. With the factor feedback, we classify these different visual representations.

4.5.1 Trigger

Controllers are used in the majority of literature for distant selection studies, as we found 118 methods using controllers as input device. One of the advantages of controllers, in contrast to head-pointing and hand-pointing methods, is that they have additional buttons for further interaction. Hence, buttons are also the most widely used tool to confirm selections in the literature in general (see Table 4.9) and especially when pointing with controllers (see Figure 4.4). In contrast to mouse input for regular 2D interaction, which is stabilized through a table, usually 3D interaction with controllers has users position the controller freely in midair. This can result in user fatigue, and more importantly regarding the confirmation of selection, can lead to hand-jitter when pressing a button [21], which in the end can lead to missing the target. To circumvent this phenomenon, many researchers had their participant confirm selections with their non-pointing hand, the offhand. Head and eye pointing methods also include head-hand pointing methods, where the direction of the ray is also dependent on the position of the hand, thus, leading to the same problem. Head-pointing methods on the other hand have the challenge of not having a direct way to confirm selections. Thus, many head-pointing techniques either have users have an additional input device with a button to confirm selection [39, 100] or rely on dwell time [19, 44, 149]. With eye-tracking methods fixations can also be used instead of a fixed dwell time [150], and also blinking [26, 98]. Furthermore, we found one method which confirmed a selection when a head ray and eye ray converged [150]. Additionally head gestures, like nodding can also be used as a confirmation tool [98, 191]. When interacting with the hands, gestures are the main form of confirming selections, where jitter can also affect the selection accuracy. Researchers used, among others, pinch gestures [26, 88, 108,

Trigger	Triggerhand	#	Citations
Button	Mainhand	120	[2, 5, 6, 10–12, 14, 24–27, 30, 32, 34, 35, 37, 38, 43, 45, 46, 48, 49, 56–58, 63, 65, 76, 78–80, 82, 83, 88, 90, 91, 96, 99, 114, 119, 123, 128, 137, 139, 141, 142, 149, 156–161, 166, 169, 170, 172, 174, 175, 179, 180, 182, 184, 186, 187, 189, 190, 194, 195, 198, 199]
	Offhand	35	[4, 8, 13, 15, 32, 39, 57, 66, 83, 100, 119, 123, 126, 131, 135–137, 149, 157]
Gesture	Mainhand	35	[7, 17, 20, 40, 49, 51, 69, 71, 88, 89, 93, 95, 108, 118, 125, 127, 129, 133, 146, 148, 168, 178, 196]
	Offhand	15	[26, 28, 39, 67, 72, 83, 141, 144, 162, 184, 190, 192, 197]
Touch	Mainhand	32	[33, 47, 52, 59, 67, 74, 84, 94, 106, 116, 128, 130, 138, 148, 155, 162, 177, 185, 193]
	Offhand	14	[28, 44, 59, 176]
Dwell	None	32	[9, 19, 26, 29, 39, 44, 85, 97, 98, 101, 102, 119, 124, 149–151, 157, 167, 181, 190, 191]
Voice	None	10	[39, 121, 143, 152, 181, 184, 198]
Other	None	14	[2, 26, 53, 68, 86, 98, 122, 150, 172, 188, 191]

Table 4.9: Results of the literature for the different trigger mechanisms we extract, further grouped by the hand that performed the selection.

125], circling with the finger [39], pointing gestures [144, 146], and poking/tapping gestures [20, 83, 89, 184]. Furthermore, with a gesture recognition system different gestures can be mapped onto different actions [20]. Using voice as trigger is a hands-free selection confirmation method, which does not lead to the Heisenberg effect. It can either be used as a clicking mechanism [39, 198], but also take semantic context of the spoken words into account and, thus, work as a disambiguation mechanism [152]. In some specific applications the usage of foot input could also be used as a trigger [53]. Fairly recently multiple works addressed crossing based methods as confirmation method [172, 191]. Crossing based confirmation methods trigger a selection, when an arbitrary line of a target gets crossed one or multiple times. No physical action is required like a button press, which can lead to hand jitter, thus, crossing techniques avoid the Heisenberg effect. However, as these techniques do not require an explicit trigger to confirm selection, the Midas touch effect can occur, especially in dense settings [173].

Dwell time is a simple method to avoid the Heisenberg effect, but leads to unnecessary waiting time until a selection is performed, which can lead to user frustration. However, if the dwell duration is set to small, the Midas touch effect can occur, and accidental selection happen [61]. Variable and user dependent dwell time can be used to mitigate this problem [167]. The results of works comparing dwell time with other trigger mechanisms like buttons or gestures conclude that button as trigger mechanism led to faster selection time in four out of five studies, especially in dense and occluded environments [39, 44, 119, 149, 190]. Dwell led to faster performances in a text input task. Gestures as trigger led to faster performance than dwell time in two out of three publications [26, 39, 190]. As button confirmation can lead to the Heisenberg effect we also labeled each method with the hand the button was pressed. However, we did not find any literature comparing one selection method with performing a selection confirmation with the offhand or mainhand. Additionally, by forcing users to confirm selection with the non-pointing hand it prevents bimanual selection, as both hands are in use to control one ray. The usage of filters to reduce jitter generally can also work to reduce hand jitter when pressing a button.

Visual	Sound	Haptic	#	Citations
✓	✓	✓	3	[175]
		✗	35	[11–15, 25, 44, 51, 71, 99, 148, 157, 172, 176]
	✗	✗	172	[4, 5, 8, 10, 19, 27–29, 32–34, 37–40, 45, 46, 48, 49, 52, 53, 56–58, 63, 65–68, 74, 76, 79, 80, 85, 86, 88–91, 93, 94, 97, 98, 100, 108, 114, 119, 121, 123, 125–127, 131, 133, 135–137, 141, 149–151, 155, 159, 160, 162, 166, 167, 169, 174, 177, 179, 180, 184–191, 193, 196, 197, 199]
		✓	4	[84, 106, 114, 118]
✗	✓	✗	15	[24, 35, 43, 96, 139, 178, 194, 198]
	✗	✗	73	[2, 6, 7, 9, 17, 20, 26, 30, 47, 59, 72, 78, 82, 83, 101, 102, 116, 119, 122, 124, 129, 130, 138, 142–144, 146, 152, 156, 158, 161, 168, 170, 181, 182, 192, 195]

Table 4.10: Results for the literature grouped by the type of feedback given after a selection.

Confirming a selection can either be implicit or explicit. Implicit trigger mechanisms often suffer from the Midas Touch problem, while explicit might suffer from the Heisenberg effect. This can be prevented by choosing a hands-free or offhand trigger mechanism. However hands-free approaches like voice trigger might not be usable in public scenarios and offhand confirmations disable bimanual interaction. Furthermore, usability and intuitiveness has to be respected when selecting a trigger mechanism.

4.5.2 Feedback

With the factor feedback, we describe the feedback mechanisms stating if a selection was correct or wrong. Feedback can be given to the user when hovering on selectable targets or after pressing the trigger. In theory, all five human senses could be used for feedback. However, only visual, audio, and haptic feedback is usable. Visual feedback is primarily used, as we found 171 methods using only visual feedback and 42 combining with auditive or haptic feedback. Auditive feedback was used standalone in 15 methods and 53 times in combination with the other feedback-modalities. Haptic feedback in total was used seven times. In Table 4.10 we present our findings on which feedback mechanisms are used in the literature.

Visual feedback in controlled studies is given mainly by coloring the selected target [174]. The color can also represent if a selection was successful or erroneous. However, for this, the selection technique needs semantic knowledge about which target is the *correct* target. Additionally, targets can also be colored when hovered over so that the user gets more visual guidance on which target would be selected on a confirmation [28, 126]. Instead of changing the color, the target could also be outlined or changed visually in any other way [38, 149, 150]. In addition to target color, also the cursor color can change when crossing a target [174]. In applications, visual feedback can come in the form of bullets [57], menus [151], or other visual interactions.

Auditive feedback is another widely used feedback method. Again, when knowledge about selectable targets exists, different sounds can be played on correct and incorrect selections. In a study about the effect of the pitch in auditory error feedback [15], the authors state that higher-pitched sounds can help users focus on their error rate, and the pitch can generally be used to give the user feedback on his performance.

Haptic feedback is the least used feedback mechanism we found in the literature. However, it has been shown that haptic feedback can guide users towards targets [1], especially in dense settings [60]. Additionally, both sound and haptic feedback can be used to guide visually impaired users [75]. Montano-Murillo et al. [114] compared slicing volume techniques with real and virtual tablets and found that users performed better when selecting targets on the physical tablet. Haptic devices can also be constrained in specific movements to help the user find targets [175]. Multiple studies also used multimodal feedback. However, the literature states that this does not affect performance either positively or negatively [175].

We found 73 methods with unknown or no feedback. On the one hand, in 2D target selection tasks, or when interacting with a touch display, feedback is not directly required as target selection is not that difficult. On the other hand, only a small amount of these unknown feedback papers explicitly stated that they would use no feedback for a specific reason. We can abbreviate that not all researchers find that feedback after a selection is a focus point of selection techniques. We also want to differentiate between the terms visualization and feedback, as in the literature, the term visual feedback often got used to describe a cursor visualization method.

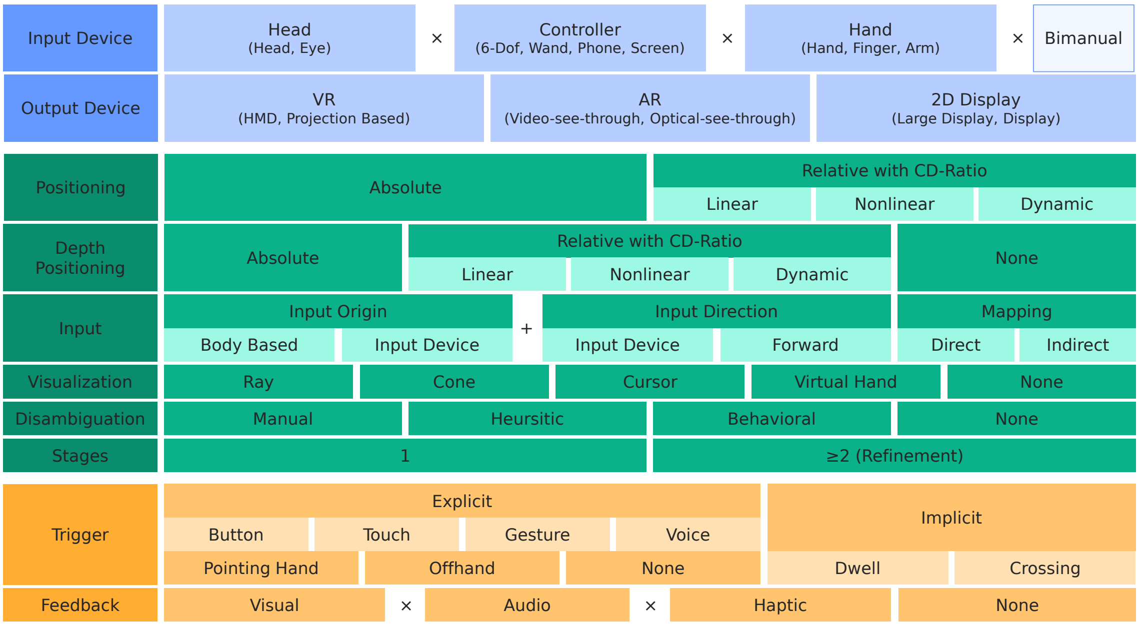
5 Design Space

Due to the wide array of different possibilities to interact with 3D environments, we propose a design space for distant selection methods. We understand that multiple challenges arise when developing a new distant selection method and aim to guide researchers to solve them all. Future work on this topic can be brought to the same level with knowledge of the problem space and the possession of a design space of options. In the following, we present our design space with 10 factors. The design space is categorized into three categories: (**D1**) Input and Output Devices, (**D2**) Selection Process, and (**D3**) Confirmation. Each category has multiple factors, with different options, gathered from analyzing the literature in Chapter 4. Figure 5.1 depicts the design space, where each row describes the levels of one factor. In the following, we explain each factor and its levels.

(D1) Input Devices
The factor input devices describes which types of devices are used as input. We propose the three levels: HEAD, CONTROLLER, and HAND, whereby the used input device should be classified into one of these categories. Suppose head movement controls the selection tool. In that case, we classify the head as an input device. However, the usage of VR-HMDs does not directly classify the head as an input device. The level CONTROLLER includes all handheld devices used for selection. The level HAND in contrast contains all selection methods using the hand, finger, or arm as a direct input modality. The levels CONTROLLER and HAND levels are different as controllers can be moved around freely and even be held at the wrist level, allowing interaction with the whole selection space. HAND based interaction does not allow this as our physical body constrains movement. This leads to a different type of interaction, more comparable with the pointing metaphor. The factor input device can be multimodal; hence multiple levels can be selected. When using *Controller* or *Hand* as input device, it should be specified if the selection technique supports or requires bimanual input. For each level, we included the mainly used input modalities in the literature we looked at.

(D1) Output Devices
We found three different levels for the factor output device, in which distant selection can be performed, namely VR, AR, and DISPLAYS. Again, each level includes multiple modalities, which can have different properties regarding selection. In this age, VR is mainly created through HMDs. However, also CAVE and stereo displays create virtual environments. Additionally, AR is also widely created with HMDs. Yet, the more accessible form of AR are video-see-through techniques via a smartphone. VR and AR mostly create 3D environments, whereby DISPLAYS visualize the environment on a 2D screen. Nevertheless, the environment can still be 3D. In the latter, we understand that distant selection is mainly interesting for large displays, but we also found some literature testing distant selection techniques for regular-sized displays.

(D2) Positioning
Positioning describes the connection between the input device and the cursor (i.e., control space and display space). This factor can be categorized into the two levels ABSOLUTE and RELATIVE.



Input Device	Head (Head, Eye) × Controller (6-Dof, Wand, Phone, Screen) × Hand (Hand, Finger, Arm) × Bimanual
Output Device	VR (HMD, Projection Based) AR (Video-see-through, Optical-see-through) 2D Display (Large Display, Display)
Positioning	Absolute Relative with CD-Ratio: Linear, Nonlinear, Dynamic
Depth Positioning	Absolute Relative with CD-Ratio: Linear, Nonlinear, Dynamic None
Input	Input Origin: Body Based, Input Device + Input Direction: Input Device, Forward Mapping: Direct, Indirect
Visualization	Ray Cone Cursor Virtual Hand None
Disambiguation	Manual Heursitic Behavioral None
Stages	1 ≥2 (Refinement)
Trigger	Explicit: Button, Touch, Gesture, Voice; Pointing Hand, Offhand, None Implicit: Dwell, Crossing
Feedback	Visual × Audio × Haptic None

Figure 5.1: Design Space for distant selection methods. Impact factors are depicted on the left, the levels of each impact factor are in the accompanying row. Levels connected with no symbols should be seen as a *XOR* connection, × indicates an *OR* connection, and + denotes an *AND* connection.

The former describes methods with a 1-to-1 CD-ratio, i.e., the connection between the motor and visual spaces matches. The latter describes methods where the input device is used to move the cursor, resulting in the CD-ratio not being 1-to-1. The CD-ratio for RELATIVE methods can usually be linear, nonlinear, or dynamic. For the large majority of methods, positioning is either ABSOLUTE or RELATIVE. However, with refinement processes, both can coexist for the same selection method. In the literature, we found that ABSOLUTE positioning techniques are used in the large majority of selection methods due to them being more immersive, being able to cover the whole display space, and reaching faster performance. In contrast, RELATIVE positioning methods reach lower error rates as they are more precise than ABSOLUTE methods.

(D2) Depth-Positioning

Depth positioning can also be enabled through ABSOLUTE or RELATIVE positioning. However, only 15% of methods in the literature implemented positioning of the cursor in the depth. Depth positioning is used to select occluded targets, as usually in 3D selection, the target closest to the user is selected. ABSOLUTE depth positioning does not enable distant occluded target selection on its own as this only covers the space reachable by the user. This can be circumvented by using an origin offset technique, which places the cursor at a defined offset in the distance, to then control it with ABSOLUTE depth positioning. RELATIVE methods are more useful for depth positioning, as a much larger space can be covered with nonlinear CD-ratios. If the selection method should be able to deal with occluded targets, depth positioning is one way to solve this. However, disambiguation and refinement techniques are also applicable.

(D2) Input
The factor input describes through which input the selection tool receives the position of the cursor. For most absolute positioning methods, an INPUT ORIGIN and INPUT DIRECTION should be specified. Most absolute positioning methods enable distant selection through raycasting, which either needs one origin and a directional vector or two points to calculate a direction. With the sublevels *body based* and *input device* we subdivide the INPUT ORIGIN even more, as it makes a difference whether the ray originates from an external handheld device or the body. Together with the sublevels *input device* and *forward* in INPUT DIRECTION, we propose every form of raycasting should be covered and differentiable. Relative positioning methods are mostly MAPPING based. The sublevels *direct* and *indirect* indicate how the input device controls the cursor. *Direct* mapping classifies techniques, where the cursor has a constant connection to the input device and movement is directly mapped onto the display. In contrast, *Indirect* mapping describes techniques, where the cursor is controlled indirectly, e.g., with a joystick or trackpad.

(D2) Visualization
The factor visualization describes the visual representation of the selection tool. We found five levels for this factor. A CURSOR in some shape got used in most techniques, whereby most authors used a spherical cursor for 3D environments. RAY and CONE visualize the ray from their input method to help users gauge the direction they are pointing at. CONE visualizations are usually combined with a disambiguation mechanism. Only a small amount of VIRTUAL HAND techniques is used for distant selection. However, its visualization can lead to higher levels of embodiment and can also be combined with other visualizations techniques. Touchscreen-based interaction mostly did not include visualization, as targets are directly selected over the screen.

(D2) Disambiguation
This factor describes whether and which disambiguation mechanism is used in a selection method. Disambiguation has the levels: MANUAL, HEURISTIC, BEHAVIORAL, and NONE. With these disambiguation mechanisms, target selection can be further extended to have the selection tool select specific targets. MANUAL disambiguation is always combined with a refinement technique, thus, needing more than one stage. HEURISTIC disambiguation selects targets in the area of the cursor, based on a heuristic. The most used heuristic is selecting the closest target to the selection tool. This improves the usability and performance of selection techniques, as distant selections can be performed with less accuracy. However, this can lead to confusion in dense environments as many targets will always be close to the cursor. BEHAVIORAL disambiguation mainly uses intent prediction of users to predict which target they want to select. Additionally, semantic input is also classified as BEHAVIORAL disambiguation. For example, users can state to select the “furthest target”. Disambiguation is often combined with splitting the selection into two stages.

(D2) Stages
The factor stages is closely connected with the factor disambiguation. Manual disambiguation mechanisms always need a second stage with a refinement technique. PROGRESSIVE REFINEMENT can be great to solve challenges from dense environments. The selection technique first selects all targets in a specific area (or based on a heuristic) to present this subset to the users and gives them a choice to select one target from this subset. However, this can lead to unnecessary additional selection time in sparse environments, as targets could also be accurately selected without a second stage. Thus, some methods propose to automatically detect if refinement is necessary, based on the density of the environment. As mentioned, we also classify techniques where absolute and relative positioning are both used sequentially as two-stage refinement techniques.

(D3) Trigger

The factor trigger has two levels: EXPLICIT and IMPLICIT. The former needs the user to explicitly select targets through an action, while the latter does not need additional user input to perform a selection. When choosing a trigger, both the Midas Touch problem and the Heisenberg effect need to be considered. The type of EXPLICIT trigger is partly dependent on the input device. We found four sublevels for EXPLICIT triggers: *button*, *touch*, *gesture*, and *voice*. The first three sublevels can further be classified if the confirmation is done with the pointing hand or the offhand. Performing an explicit trigger action with the pointing hand can cause jitter, as the Heisenberg effect states. Thus, many selections methods circumvented this by having users confirm a selection with their other hand. We found two sublevels for implicit selection: *dwell* and *crossing*. Both of these can lead to the Midas Touch problem but solve the Heisenberg effect. In the literature, the large majority of researchers settled for an explicit trigger, giving the users more control over their actions.

(D3) Feedback

The last factor describes the feedback after a performed selection. With the levels VISUAL, AUDIO, and HAPTIC we cover all types of feedback used in the literature we analyzed. It is also possible to give no feedback after a performed selection. VISUAL feedback is used in the majority of selection techniques and ranges from colored targets, over colored cursors, up to contextual visual events. AUDIO feedback is also frequently used to emulate a button click or to give users different sounds depending on if the selection was correct or erroneous. We found that HAPTIC feedback was used only in a very small number of publications. AUDIO and HAPTIC feedback can also be used to guide visually impaired users towards targets.

6 Discussion

In the previous sections, we laid out our process of analyzing the literature and then created a design space based on our findings. The goal of this work was to get a deeper understanding of the different possibilities to create a distant selection technique. We can describe the properties of selection methods with our proposed design factors and compare them based on these categories. Thus, we argue that our design space is a step towards finding a unified solution for distant selection.

In total, we extracted 302 distant selection methods from 146 papers using the PRISMA guidelines [113]. Using additional coding, we then extracted ten impact factors that make up three categories (D1-D3). This allows to clearly categorize distant selection methods when excluding feedback as it is a multimodal impact factor. This is apparent from the 175 different unique methods, which we could extract from the 302 methods using our design space first level (see Figure 6.1). On the other hand, this is only a fraction of the total option space of more than 21,000 options. We acknowledge that researchers have narrowed the option space down to a small subset. Nevertheless, we argue that this subset is still too plentiful to be managed without guidance. Thus, our investigation itself showed an inherent need for more structure when designing distant selection methods. With our design space we aim to give that guidance, as researchers can use the knowledge provided by this work to pinpoint solutions for emergent challenges.

In literature, we found two main challenges, which are playing against each other: user immersion and fatigue versus dense and occluded environments. When looking at the implementation of distant target selection in current consumer VR-HMDs, we see that they exclusively rely on very basic directional raycasting, usually with a 6-DoF controller. Consumers today primarily use VR for the purpose of entertainment. For these types of virtual experiences, immersion is one of the key factors for a satisfactory user experience [117]. Immersion quantifies all factors that play a role in the technical realization of making a virtual environment match a real environment [154]. This includes the quality of graphics, self-location through good body tracking, and also interaction, that is, in our case, selection. Thus, how we select in VR affects the level of perceivable immersion. The literature suggests that users find selection with a 6-DoF controller often straightforward to learn, and only little fatigue occurs. A controller, in contrast to head-hand pointing, bypasses the Gorilla Arm Syndrome [50] as it can be held close to the wrist. This form of distant target pointing is also intuitive for most users, as it is very similar to pointing with a laser-pointer. However, forward-raycasting in its basic form is unable to select occluded targets and has difficulties in dealing with densely populated environments. For these cases, the literature proposed multiple different solutions. VR, although mainly, is not only used to serve the entertainment industry and build fully immersive environments. Looking at data in a virtual environment can be profitable, as viewing graphs in three dimensions can be helpful to deal with visual clutter and the general trend of data becoming increasingly high dimensional. These environments usually are densely populated and introduce many occluded targets. Especially refinement strategies are applicable to dense environments. Furthermore, occluded targets can be selected with different depth selection strategies. However, these techniques mostly add a substantial amount of complexity and are

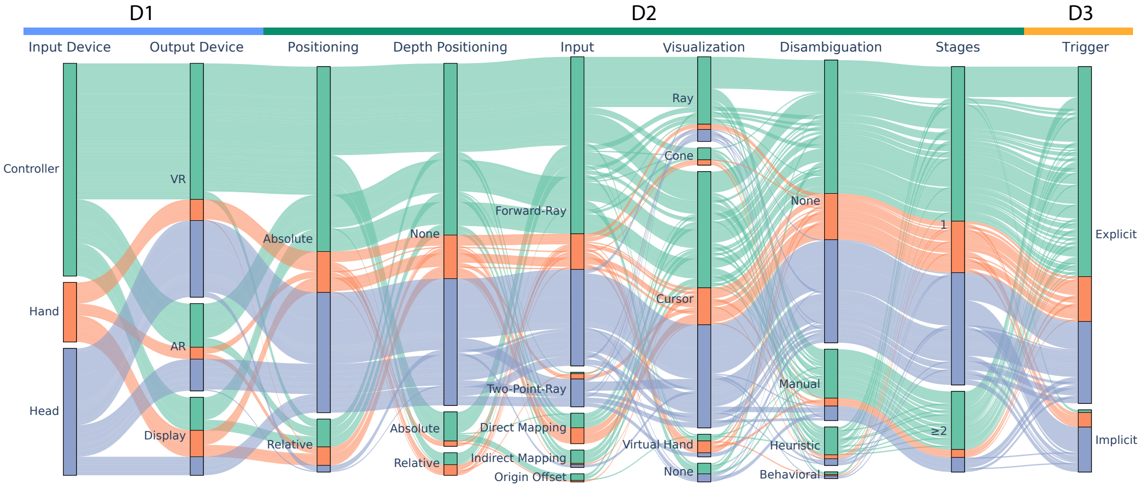


Figure 6.1: The literature categorized for each impact factor. Each line represents one method.

also often designed specifically for more complex environments. To circumvent unnecessary additional refinement steps in sparse environments, a selection technique can automatically detect if refinement is needed and automatically switch between an easy and a more sophisticated solution. Disambiguation factors can also improve basic raycasting in sparse settings and improve the accuracy of absolute pointing strategies. Currently, heuristic approaches are the most used non-refinement disambiguation mechanism. However, with the rising research interest in intention prediction, behavioral disambiguation mechanisms might also be used in the future. Behavioral disambiguation might also solve the challenge of heuristic disambiguation in dense environments, of switching between close targets, and make it more user-friendly.

When trying to unify distant selection, developers always have to weigh up performance and usability. One example is the Heisenberg effect. Although there are clear indications that performing an explicit selection with the pointing hand leads to hand-jitter, most techniques still use the pointing hand to confirm the selection, as this is much more usable for users. Additionally, confirming with the non-pointing hand also restricts the use of bimanual selection. This predicament can be translated into many other factors of our design space. Relative positioning has often been shown to be more precise, and absolute positioning leads to more immersion and decreased selection times. Nevertheless, when evaluating our design space with the literature, we also find gaps in research in certain areas. In some cases, these gaps might be explainable due the technical progression of hardware, and some factor-combinations being impossible to implement. However, we suggest a need for more fine granular studies regarding the effect of different levels in one factor. Researchers should also follow design guidelines on how to design studies to evaluate 3D interaction techniques [18]. A unified selection task would allow the comparison of different methods between studies, and not only within. With a unified selection task and concrete strategies through our design space, we can determine which selection strategy is best at being immersive, easy to learn for users, and selecting targets in dense and occluded environments.

7 Conclusion

In this work, we analyzed existing work on distant selection methods to build a design space. We collected a corpus of selection techniques by following the PRISMA guidelines [113]. Through a systematic comparison and clustering of this literature, we were able to propose ten impact factors grouped into three categories. We found that different combinations of output and input devices are used more frequently. Furthermore, different input devices lead to diverging selection metaphors, resulting in varying amounts of immersion and fatigue for the user. Whereas hands-free interaction might be desirable for AR applications, controllers offer more flexibility due to their additional DoF and handheld capability and thus might prevail for VR. Moreover, different methods vary in their resemblance to real-world interaction, therefore achieving varying levels of immersion. However, independent of the chosen metaphors, we found that many proposed selection techniques often brought up application-driven solutions. During our analysis, we found that being able to interact with different kinds of environments while still maintaining usability is currently one of the main challenges to solve. User fatigue and usability should always be highly prioritized to enable quick and accurate interaction with immersive environments. However, the same selection technique should also be able to deal with high object density and occluded objects. With our design space, the process of comparing methods and creating one single unified method can be streamlined. However, it is impossible to compare every single factor-combination from our design space due to the option-space encompassing 21,000 combinations. This suggests an inherent need for a more structural approach to solving this challenge.

We propose that our design space enables and simplifies future work on the topic of distant selection methods. In the future, studies comparing selection methods have access to our proposed factors to assist them in their study design. While we identified already vastly different selections methods, the design space itself opens up the opportunity for even more exploration. Thus, we enable designers to explore even more possibilities, eventually helping users to optimally interact with the environment. We should strive to systematically find the best solution, rather than compromising on a technique that just works. Due to the size of the design space, it is crucial to understand the influence of the impact factors in more detail. Thus, it will be important to study impact factors independently allowing for a better understanding towards a optimal design. Future work should therefore build their studies around one factor to further reveal the advantages and disadvantages of each factor's levels. At the very least, future work profits from this design space by gaining an overview of the multiple options how distant selection in VR can be constructed. But we argue that in the long run our design space will bring transparency to the currently scattered landscape of different distant selection methods. Ultimately, we hope that we can bring clarity to the field and help to converge to an universal selection concept for distant selection.

Bibliography

[1] Teemu Tuomas Ahmaniemi, Vuokko Tuulikki Lantz. "Augmented Reality Target Finding Based on Tactile Cues". In: *Proceedings of the 2009 International Conference on Multimodal Interfaces*. ICMI-MLMI '09. Cambridge, Massachusetts, USA: Association for Computing Machinery, 2009, pp. 335–342. ISBN: 9781605587721. DOI: `10.1145/1647314.1647383` (cit. on p. 46).

[2] Carlos Andujar, Ferran Argelaguet. "Anisomorphic ray-casting manipulation for interacting with 2D GUIs". In: *Computers & Graphics* 31.1 (Jan. 2007), pp. 15–25. DOI: `10.1016/j.cag.2006.09.003` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 42, 44, 45).

[3] Ferran Argelaguet, Carlos Andujar. "A survey of 3D object selection techniques for virtual environments". In: *Computers & Graphics* 37.3 (2013), pp. 121–136. ISSN: 0097-8493. DOI: `10.1016/j.cag.2012.12.003` (cit. on pp. 13, 16, 17, 21, 39).

[4] Ferran Argelaguet, Carlos Andujar. "Efficient 3D Pointing Selection in Cluttered Virtual Environments". In: *IEEE Computer Graphics and Applications* 29.6 (Nov. 2009), pp. 34–43. ISSN: 1558-1756. DOI: `10.1109/MCG.2009.117` (cit. on pp. 27, 30, 31, 34, 36–38, 44, 45).

[5] Ferran Argelaguet, Carlos Andujar. "Visual Feedback Techniques for Virtual Pointing on Stereoscopic Displays". In: *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*. VRST '09. Kyoto, Japan: Association for Computing Machinery, 2009, pp. 163–170. ISBN: 9781605588698. DOI: `10.1145/1643928.1643966` (cit. on pp. 27, 30, 31, 34, 36, 38, 39, 41, 42, 44, 45).

[6] Ferran Argelaguet, Carlos Andujar, Ramon Trueba. "Overcoming Eye-Hand Visibility Mismatch in 3D Pointing Selection". In: *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology*. VRST '08. Bordeaux, France: Association for Computing Machinery, 2008, pp. 43–46. ISBN: 9781595939517. DOI: `10.1145/1450579.1450588` (cit. on pp. 27, 30, 31, 34, 36–38, 44, 45).

[7] Cagan Arslan, Yosra Rekik, Laurent Grisoni. "E-Pad: Large Display Pointing in a Continuous Interaction Space around a Mobile Device". In: *Proceedings of the 2019 on Designing Interactive Systems Conference*. DIS '19. San Diego, CA, USA: Association for Computing Machinery, 2019, pp. 1101–1108. ISBN: 9781450358507. DOI: `10.1145/3322276.3322284` (cit. on pp. 27, 30, 31, 34, 36–38, 44, 45).

[8] K. Asai, N. Osawa, H. Takahashi, Y.Y. Sugimoto, S. Yamazaki, A. Samejima, T. Tanimae. "Eye mark pointer in immersive projection display". In: *Proceedings IEEE Virtual Reality 2000*. VR '00. IEEE, Mar. 2000, pp. 125–132. DOI: `10.1109/VR.2000.840490` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[9] Sabarish Babu, Ming-Han Tsai, Ting-Wei Hsu, Jung-Hong Chuang. “An Evaluation of the Efficiency of Popular Personal Space Pointing versus Controller Based Spatial Selection in VR”. In: *ACM Symposium on Applied Perception 2020*. SAP ’20. Virtual Event, USA: Association for Computing Machinery, 2020. ISBN: 9781450376181. DOI: `10.1145/3385955.3407939` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[10] Marc Baloup, Thomas Pietrzak, Géry Casiez. “RayCursor: A 3D Pointing Facilitation Technique Based on Raycasting”. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI ’19. Glasgow, Scotland UK: Association for Computing Machinery, 2019, pp. 1–12. ISBN: 9781450359702. DOI: `10.1145/3290605.3300331` (cit. on pp. 27, 30, 31, 34, 36, 38, 40–42, 44, 45).

[11] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Duc Minh Pham, Wolfgang Stuerzlinger. “Do Head-Mounted Display Stereo Deficiencies Affect 3D Pointing Tasks in AR and VR?” In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces*. VR ’19. IEEE, Mar. 2019, pp. 585–592. DOI: `10.1109/VR.2019.8797975` (cit. on pp. 27, 29–31, 34, 36, 38, 44, 45).

[12] Anil Ufuk Batmaz, Aunnoy K Mutasim, Wolfgang Stuerzlinger. “Precision vs. Power Grip: A Comparison of Pen Grip Styles for Selection in Virtual Reality”. In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. VRW ’20. IEEE, Mar. 2020, pp. 23–28. DOI: `10.1109/VRW50115.2020.00012` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[13] Anil Ufuk Batmaz, Wolfgang Stuerzlinger. “Effect of Fixed and Infinite Ray Length on Distal 3D Pointing in Virtual Reality”. In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI EA ’20. Honolulu, HI, USA: Association for Computing Machinery, 2020, pp. 1–10. ISBN: 9781450368193. DOI: `10.1145/3334480.3382796` (cit. on pp. 27, 30, 31, 34, 36–38, 41, 42, 44, 45).

[14] Anil Ufuk Batmaz, Wolfgang Stuerzlinger. “Effects of 3D Rotational Jitter and Selection Methods on 3D Pointing Tasks”. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. VR ’19. IEEE, Mar. 2019, pp. 1687–1692. DOI: `10.1109/VR.2019.8798038` (cit. on pp. 27, 30, 31, 34, 36, 38, 41, 42, 44, 45).

[15] Anil Ufuk Batmaz, Wolfgang Stuerzlinger. “The Effect of Pitch in Auditory Error Feedback for Fitts’ Tasks in Virtual Reality Training Systems”. In: *2021 IEEE Virtual Reality and 3D User Interfaces*. VR ’21. IEEE, Mar. 2021, pp. 85–94. DOI: `10.1109/VR50410.2021.00029` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[16] Amy L. Baylor. “Promoting motivation with virtual agents and avatars: role of visual presence and appearance”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 364.1535 (2009), pp. 3559–3565. DOI: `10.1098/rstb.2009.0148` (cit. on p. 35).

[17] Abdelkader Bellarbi, Nadia Zenati, Samir Otmane, Hayet Belghit, Samir Benbelkacem, Assia Messaci, Mahfoud Hamidia. “A 3D interaction technique for selection and manipulation distant objects in augmented reality”. In: *2017 5th International Conference on Electrical Engineering - Boumerdes*. ICEE-B ’17. IEEE, Oct. 2017, pp. 1–5. DOI: `10.1109/ICEE-B.2017.8192012` (cit. on pp. 27, 30, 31, 34, 36, 38, 40–42, 44, 45).

[18] Joanna Bergström, Tor-Salve Dalsgaard, Jason Alexander, Kasper Hornbæk. “How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI ’21. New York, NY, USA: Association for Computing Machinery, 2021. ISBN: 9781450380966 (cit. on pp. 14, 17, 52).

[19] Jonas Blattgerste, Patrick Renner, Thies Pfeiffer. “Advantages of Eye-Gaze over Head-Gaze-Based Selection in Virtual and Augmented Reality under Varying Field of Views”. In: *Proceedings of the Workshop on Communication by Gaze Interaction*. COGAIN ’18. Warsaw, Poland: Association for Computing Machinery, 2018. ISBN: 9781450357906. DOI: 10.1145/3206343.3206349 (cit. on pp. 27, 30–32, 34, 36, 38, 43–45).

[20] Thanatorn Boonnak, Vasaka Visoottiviseth, Jason Haga, Dylan Kobayashi, Jason Leigh. “Integration of Gesture Control with Large Display Environments Using SAGE2”. In: *2018 15th International Joint Conference on Computer Science and Software Engineering*. JCSSE ’18. IEEE, July 2018, pp. 1–5. DOI: 10.1109/JCSSE.2018.8457323 (cit. on pp. 27, 30, 31, 34, 36–38, 44, 45).

[21] D. Bowman, Chadwick A. Wingrave, J. B. Campbell, V. Q. Ly. “Using Pinch Gloves(TM) for both Natural and Abstract Interaction Techniques in Virtual Environments”. In: 2001 (cit. on pp. 17, 43).

[22] Doug A Bowman, Donald B Johnson, Larry F Hodges. “Testbed evaluation of virtual environment interaction techniques”. In: *Presence* 10.1 (2001), pp. 75–95 (cit. on p. 16).

[23] Wolfgang Büschel, Anke Lehmann, Raimund Dachselt. “MIRIA: A Mixed Reality Toolkit for the In-Situ Visualization and Analysis of Spatio-Temporal Interaction Data”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI ’21. Yokohama, Japan: ACM, May 2021. ISBN: 978-1-4503-8096-6/21/05. DOI: 10.1145/3411764.3445651 (cit. on p. 18).

[24] Jeffrey Cashion, Chadwick Wingrave, Joseph J. LaViola. “Optimal 3D selection technique assignment using real-time contextual analysis”. In: *2013 IEEE Symposium on 3D User Interfaces*. 3DUI ’13. IEEE, Mar. 2013, pp. 107–110. DOI: 10.1109/3DUI.2013.6550205 (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 41, 44, 45).

[25] Jeffrey Cashion, Chadwick Wingrave, Joseph J. LaViola Jr. “Dense and Dynamic 3D Selection for Game-Based Virtual Environments”. In: *IEEE Transactions on Visualization and Computer Graphics*. TVCG ’12 18.4 (Apr. 2012), pp. 634–642. ISSN: 1941-0506. DOI: 10.1109/TVCG.2012.40 (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 41, 44, 45).

[26] Ishan Chatterjee, Robert Xiao, Chris Harrison. “Gaze+Gesture: Expressive, Precise and Targeted Free-Space Interactions”. In: *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ICMI ’15. Seattle, Washington, USA: Association for Computing Machinery, 2015, pp. 131–138. ISBN: 9781450339124. DOI: 10.1145/2818346.2820752 (cit. on pp. 27, 30, 31, 34–38, 40, 43–45).

[27] Sibo Chen, Junce Wang, Santiago Guerra, Neha Mittal, Soravis Prakkamakul. “Exploring Word-Gesture Text Entry Techniques in Virtual Reality”. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI EA ’19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–6. ISBN: 9781450359719. DOI: 10.1145/3290607.3312762 (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[28] Yuan Chen, Keiko Katsuragawa, Edward Lank. “Understanding Viewport- and World-Based Pointing with Everyday Smart Devices in Immersive Augmented Reality”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI ’20. Honolulu, HI, USA: Association for Computing Machinery, 2020, pp. 1–13. ISBN: 9781450367080. DOI: `10.1145/3313831.3376592` (cit. on pp. 27, 30–38, 40, 44, 45).

[29] Kelvin Cheng, Masahiro Takatsuka. “Initial Evaluation of a Bare-Hand Interaction Technique for Large Displays Using a Webcam”. In: *Proceedings of the 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. EICS ’09. Pittsburgh, PA, USA: Association for Computing Machinery, 2009, pp. 291–296. ISBN: 9781605586007. DOI: `10.1145/1570433.1570487` (cit. on pp. 27, 30, 31, 34, 36–38, 44, 45).

[30] Logan D. Clark, Aakash B. Bhagat, Sara L. Riggs. “Extending Fitts’ law in three-dimensional virtual environments with current low-cost virtual reality technology”. In: *International Journal of Human-Computer Studies* 139 (July 2020), p. 102413. DOI: `10.1016/j.ijhcs.2020.102413` (cit. on pp. 25, 27, 30, 31, 34, 36, 38, 44, 45).

[31] C Comeau. “Headsight television system provides remote surveillance”. In: *Electronics* (1961), pp. 86–90 (cit. on p. 15).

[32] Nathan Cournia, John D. Smith, Andrew T. Duchowski. “Gaze- vs. Hand-Based Pointing in Virtual Environments”. In: *CHI ’03 Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’03. Ft. Lauderdale, Florida, USA: Association for Computing Machinery, 2003, pp. 772–773. ISBN: 1581136374. DOI: `10.1145/765891.765982` (cit. on pp. 27, 30–32, 34, 36, 38, 44, 45).

[33] Rajkumar Darbar, Joan Odicio-Vilchez, Thibault Lainé, Arnaud Prouzeau, Martin Hachet. “Text Selection in AR-HMD Using a Smartphone as an Input Device”. In: *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. VRW ’21. IEEE, Mar. 2021, pp. 526–527. DOI: `10.1109/VRW52623.2021.00145` (cit. on pp. 26, 27, 29–31, 34–36, 38, 44, 45).

[34] Kaushik Das, Christoph W. Borst. “An evaluation of menu properties and pointing techniques in a projection-based VR environment”. In: *2010 IEEE Symposium on 3D User Interfaces*. 3DUI ’10. IEEE, Mar. 2010, pp. 47–50. DOI: `10.1109/3DUI.2010.5444721` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[35] J. De Boeck, T. De Weyer, C. Raymaekers, K. Coninx. “Using the Non-Dominant Hand for Selection in 3D”. In: *3D User Interfaces (3DUI’06)*. 3DUI ’06. IEEE, Mar. 2006, pp. 53–58. DOI: `10.1109/VR.2006.140` (cit. on pp. 21, 27, 30–32, 34, 36, 38, 44, 45).

[36] Tara Donker, Ilja Cornelisz, Chris van Klaveren, Annemieke van Straten, Per Carlbring, Pim Cuijpers, Jean-Louis van Gelder. “Effectiveness of Self-guided App-Based Virtual Reality Cognitive Behavior Therapy for Acrophobia: A Randomized Clinical Trial”. In: *JAMA Psychiatry* 76.7 (July 2019), pp. 682–690. ISSN: 2168-622X. DOI: `10.1001/jamapsychiatry.2019.0219` (cit. on p. 18).

[37] Andrew Dunk, Adrian Haffegee, Vassil Alexandrov. “Selection methods for interactive creation and management of objects in 3D immersive environments”. In: *Procedia Computer Science* 1.1 (May 2010), pp. 2609–2617. DOI: `10.1016/j.procs.2010.04.294` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[38] Thierry Duval, Cedric Fleury. “An Asymmetric 2D Pointer/3D Ray for 3D Interaction within Collaborative Virtual Environments”. In: *Proceedings of the 14th International Conference on 3D Web Technology*. Web3D ’09. Darmstadt, Germany: Association for Computing Machinery, 2009, pp. 33–41. ISBN: 9781605584324. DOI: `10.1145/1559764.1559769` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[39] Augusto Esteves, Yonghwan Shin, Ian Oakley. “Comparing selection mechanisms for gaze input techniques in head-mounted displays”. In: *International Journal of Human-Computer Studies* 139 (July 2020), p. 102414. DOI: `10.1016/j.ijhcs.2020.102414` (cit. on pp. 27, 30, 31, 34, 36, 38, 43–45).

[40] Tiare Feuchtner, Jörg Müller. “Extending the Body for Interaction with Reality”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17. Denver, Colorado, USA: Association for Computing Machinery, 2017, pp. 5145–5157. ISBN: 9781450346559. DOI: `10.1145/3025453.3025689` (cit. on pp. 27, 30, 31, 34, 36, 38, 39, 41, 44, 45).

[41] Paul M Fitts. “The information capacity of the human motor system in controlling the amplitude of movement.” In: *Journal of experimental psychology* 47.6 (1954), p. 381. DOI: `10.1037/h0055392` (cit. on p. 17).

[42] W. Freeman, C. D. Weissman. “Television control by hand gestures”. In: 1994 (cit. on p. 15).

[43] Scott Frees, G. Drew Kessler, Edwin Kay. “PRISM Interaction for Enhancing Control in Immersive Virtual Environments”. In: *ACM Trans. Comput.-Hum. Interact.* 14.1 (May 2007), 2–es. ISSN: 1073-0516. DOI: `10.1145/1229855.1229857` (cit. on pp. 27, 30, 31, 34, 36, 38, 41, 44, 45).

[44] Priya Ganapathi, Keyur Sorathia. “Investigating Controller Less Input Methods for Smartphone Based Virtual Reality Platforms”. In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. MobileHCI ’18. Barcelona, Spain: Association for Computing Machinery, 2018, pp. 166–173. ISBN: 9781450359412. DOI: `10.1145/3236112.3236136` (cit. on pp. 27, 30, 31, 34, 36, 38, 43–45).

[45] Jérôme Grosjean, Jérôme Simonin, Eric Galin, Stéphane Mérillou. “3D Arrow: A virtual pointer for immersive sculpting”. In: *2011 IEEE Virtual Reality Conference*. VR ’11. IEEE, Mar. 2011, pp. 209–210. DOI: `10.1109/VR.2011.5759472` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 41, 44, 45).

[46] Tovi Grossman, Ravin Balakrishnan. “The Design and Evaluation of Selection Techniques for 3D Volumetric Displays”. In: *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology*. UIST ’06. Montreux, Switzerland: Association for Computing Machinery, 2006, pp. 3–12. ISBN: 1595933131. DOI: `10.1145/1166253.1166257` (cit. on pp. 27, 30, 31, 34, 36, 38, 40–42, 44, 45).

[47] Jan Gugenheimer, David Dobbelstein, Christian Winkler, Gabriel Haas, Enrico Rukzio. “FaceTouch: Enabling Touch Interaction in Display Fixed UIs for Mobile Virtual Reality”. In: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. UIST ’16. Tokyo, Japan: Association for Computing Machinery, 2016, pp. 49–60. ISBN: 9781450341899. DOI: `10.1145/2984511.2984576` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[48] Gerwin de Haan, Michal Koutek, Frits H. Post. “IntenSelect: Using Dynamic Object Rating for Assisting 3D Object Selection”. In: *Eurographics Symposium on Virtual Environments*. Ed. by Erik Kjems, Roland Blach. The Eurographics Association, 2005. ISBN: 978-3-905674-06-4. DOI: `10.2312/EGVE/IPT_EGVE2005/201-209` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[49] Pantea Habibi, Debaleena Chattopadhyay. “The impact of handedness on user performance in touchless input”. In: *International Journal of Human-Computer Studies* 149 (2021), p. 102600. ISSN: 1071-5819. DOI: `10.1016/j.ijhcs.2021.102600` (cit. on pp. 27, 30–32, 34, 36, 38, 44, 45).

[50] Jeffrey T. Hansberger, Chao Peng, Shannon L. Mathis, Vaidyanath Areyur Shanthakumar, Sarah C. Meacham, Lizhou Cao, Victoria R. Blakely. “Dispelling the Gorilla Arm Syndrome: The Viability of Prolonged Gesture Interactions”. In: *Virtual, Augmented and Mixed Reality*. Ed. by Stephanie Lackey, Jessie Chen. Cham: Springer International Publishing, 2017, pp. 505–520. ISBN: 978-3-319-57987-0 (cit. on p. 51).

[51] Faizan Haque, Mathieu Nancel, Daniel Vogel. “Myopoint: Pointing and Clicking Using Forearm Mounted Electromyography and Inertial Motion Sensors”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI ’15. Seoul, Republic of Korea: Association for Computing Machinery, 2015, pp. 3653–3656. ISBN: 9781450331456. DOI: `10.1145/2702123.2702133` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[52] Jeremy Hartmann, Daniel Vogel. “An Evaluation of Mobile Phone Pointing in Spatial Augmented Reality”. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA ’18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–6. ISBN: 9781450356213. DOI: `10.1145/3170427.3188535` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[53] Benjamin Hatscher, Maria Luz, Lennart E. Nacke, Norbert Elkmann, Veit Müller, Christian Hansen. “GazeTap: Towards Hands-Free Interaction in the Operating Room”. In: *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. ICMI ’17. Glasgow, UK: Association for Computing Machinery, 2017, pp. 243–251. ISBN: 9781450355438. DOI: `10.1145/3136755.3136759` (cit. on pp. 18, 27, 30, 31, 34, 36, 38, 44, 45).

[54] Carrie Heeter. “Being There: The Subjective Experience of Presence”. In: *Presence: Teleoperators and Virtual Environments* 1.2 (May 1992), pp. 262–271. DOI: `10.1162/pres.1992.1.2.262`. eprint: `https://direct.mit.edu/pvar/article-pdf/1/2/262/1622433/pres.1992.1.2.262.pdf` (cit. on p. 33).

[55] ML Heilig. *Sensorama Simulator. USA Patent 3050870*. 1962 (cit. on p. 15).

[56] Rorik Henrikson, Tovi Grossman, Sean Trowbridge, Daniel Wigdor, Hrvoje Benko. “Head-Coupled Kinematic Template Matching: A Prediction Model for Ray Pointing in VR”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI ’20. Honolulu, HI, USA: Association for Computing Machinery, 2020, pp. 1–14. ISBN: 9781450367080. DOI: `10.1145/3313831.3376489` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[57] Katharina Anna Maria Heydn, Marc Philipp Dietrich, Marcus Barkowsky, Götz Winterfeldt, Sebastian von Mammen, Andreas Nüchter. “The Golden Bullet: A Comparative Study for Target Acquisition, Pointing and Shooting”. In: *2019 11th International Conference on Virtual Worlds and Games for Serious Applications*. VS-Games ’19. IEEE, Sept. 2019, pp. 1–8. DOI: `10.1109/VS-Games.2019.8864589` (cit. on pp. 27, 30–32, 34, 36, 38, 44, 45).

[58] Juan David Hincapié-Ramos, Kasim Ozacar, Pourang P. Irani, Yoshifumi Kitamura. “GyroWand: IMU-Based Raycasting for Augmented Reality Head-Mounted Displays”. In: *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*. SUI ’15. Los Angeles, California, USA: Association for Computing Machinery, 2015, pp. 89–98. ISBN: 9781450337038. DOI: `10.1145/2788940.2788947` (cit. on pp. 27, 30, 31, 34, 36–38, 44, 45).

[59] Teresa Hirzle, Jan Rixen, Jan Gugenheimer, Enrico Rukzio. “WatchVR: Exploring the Usage of a Smartwatch for Interaction in Mobile Virtual Reality”. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA ’18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–6. ISBN: 9781450356213. DOI: `10.1145/3170427.3188629` (cit. on pp. 27, 30, 31, 34–38, 44, 45).

[60] Faustina Hwang, Simeon Keates, Pat Langdon, P. Clarkson. “Multiple haptic targets for motion-impaired computer users”. In: Jan. 2003, pp. 41–48. DOI: `10.1145/642611.642620` (cit. on p. 46).

[61] Toshiya Isomoto, Toshiyuki Ando, Buntarou Shizuki, Shin Takahashi. “Dwell Time Reduction Technique Using Fitts’ Law for Gaze-Based Target Acquisition”. In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. ETRA ’18. Warsaw, Poland: Association for Computing Machinery, 2018. ISBN: 9781450357067. DOI: `10.1145/3204493.3204532` (cit. on p. 44).

[62] Santiago González Izard, Juan A Juanes, Francisco J García Peñalvo, Jesús Mª Gonçalvez Estella, Mª José Sánchez Ledesma, Pablo Ruisoto. “Virtual reality as an educational and training tool for medicine”. In: *Journal of medical systems* 42.3 (2018), pp. 1–5. DOI: `10.1007/s10916-018-0900-2` (cit. on p. 18).

[63] Bret Jackson, Logan B Caraco, Zahara M Spilka. “Arc-Type and Tilt-Type: Pen-Based Immersive Text Input for Room-Scale VR”. In: *Symposium on Spatial User Interaction*. SUI ’20. Virtual Event, Canada: Association for Computing Machinery, 2020. ISBN: 9781450379434. DOI: `10.1145/3385959.3418454` (cit. on pp. 26, 27, 30, 31, 34, 36, 38, 44, 45).

[64] Robert J. K. Jacob. “What You Look at is What You Get: Eye Movement-Based Interaction Techniques”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’90. Seattle, Washington, USA: Association for Computing Machinery, 1990, pp. 11–18. ISBN: 0201509326. DOI: `10.1145/97243.97246` (cit. on pp. 17, 31).

[65] Izabelle Janzen, Vasanth K. Rajendran, Kellogg S. Booth. “Modeling the Impact of Depth on Pointing Performance”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. San Jose, California, USA: Association for Computing Machinery, 2016, pp. 188–199. ISBN: 9781450333627. DOI: `10.1145/2858036.2858244` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[66] Ricardo Jota, Miguel A. Nacenta, J. Jorge, M. Carpendale, S. Greenberg. “A comparison of ray pointing techniques for very large displays”. In: *Graphics Interface*. 2010. DOI: `10.11575/PRISM/30630` (cit. on pp. 21, 27, 30, 31, 34, 36–38, 44, 45).

[67] Whie Jung, Woojin Cho, Hayun Kim, Woontack Woo. "BoostHand : Distance-free Object Manipulation System with Switchable Non-linear Mapping for Augmented Reality Classrooms". In: *2017 IEEE International Symposium on Mixed and Augmented Reality*. ISMAR Adjunct '17. IEEE, Oct. 2017, pp. 321–325. DOI: `10.1109/ISMAR-Adjunct.2017.96` (cit. on pp. 27, 30, 31, 34, 36, 38, 41, 44, 45).

[68] P. Kanjanasupark, N. Chotikakamthorn. "A Light-Pointing Method for Virtual Object Selection and Manioulation in A Monoscopic Viewing System". In: *2005 5th International Conference on Information Communications Signal Processing*. ICICS '05'. IEEE, Dec. 2005, pp. 1076–1079. DOI: `10.1109/ICICS.2005.1689218` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[69] Keiko Katsuragawa, Krzysztof Pietroszek, James R. Wallace, Edward Lank. "Watchpoint: Freehand Pointing with a Smartwatch in a Ubiquitous Display Environment". In: *Proceedings of the International Working Conference on Advanced Visual Interfaces*. AVI '16. Bari, Italy: Association for Computing Machinery, 2016, pp. 128–135. ISBN: 9781450341318. DOI: `10.1145/2909132.2909263` (cit. on pp. 37, 44).

[70] Rimsha Khan, Farooque Azam, Sheeraz Ahmed, Waseem Anwar, Rudeema Chughtai, Anum Farid. "Comparative Analysis of Interaction Techniques in Virtual Reality". In: *2020 IEEE 23rd International Multitopic Conference (INMIC)*. INMIC '20. IEEE, 2020, pp. 1–6. DOI: `10.1109/INMIC50486.2020.9318194` (cit. on p. 13).

[71] Daniel Kharlamov, Brandon Woodard, Liudmila Tahai, Krzysztof Pietroszek. "TickTockRay: Smartwatch-Based 3D Pointing for Smartphone-Based Virtual Reality". In: *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*. VRST '16. Munich, Germany: Association for Computing Machinery, 2016, pp. 365–366. ISBN: 9781450344913. DOI: `10.1145/2993369.2996311` (cit. on pp. 27, 30, 31, 34, 36–38, 44, 45).

[72] Kiana Kia, Jaejin Hwang, In-Sop Kim, Hakim Ishak, Jeong Ho Kim. "The effects of target size and error rate on the cognitive demand and stress during augmented reality interactions". In: *Applied Ergonomics* 97 (2021), p. 103502. ISSN: 0003-6870. DOI: `10.1016/j.apergo.2021.103502` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[73] Konstantina Kilteni, Raphaela Groten, Mel Slater. "The Sense of Embodiment in Virtual Reality". In: *Presence: Teleoperators and Virtual Environments* 21.4 (Nov. 2012), pp. 373–387. ISSN: 15313263. DOI: `10.1162/PRES_a_00124` (cit. on p. 35).

[74] Daehwan Kim, Hye-mi Kim, Hang-Kee Kim, Su-Ran Park, Ki-Suk Lee, Ki-Hong Kim. "ThunderPunch: A bare-hand, gesture-based, large interactive display interface with upper-body-part detection in a top view". In: *IEEE Computer Graphics and Applications* 38.5 (Sept. 2018), pp. 100–111. ISSN: 1558-1756. DOI: `10.1109/MCG.2018.053491734` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[75] Kibum Kim, Xiangshi Ren, Seungmoon Choi, Hong Z. Tan. "Assisting people with visual impairments in aiming at a target on a large wall-mounted display". In: *International Journal of Human-Computer Studies* 86 (2016), pp. 109–120. ISSN: 1071-5819. DOI: `10.1016/j.ijhcs.2015.10.002` (cit. on p. 46).

[76] Woojoo Kim, Junyoung Jung, Shuping Xiong. "VRMouse: Mouse Emulation with the VR Controller for 2D Selection in VR". In: *Advances in Intelligent Systems and Computing*. Springer International Publishing, 2020, pp. 663–670. DOI: `10.1007/978-3-030-51828-8_87` (cit. on pp. 27, 30, 31, 34–38, 44, 45).

[77] Masatomo Kobayashi, Takeo Igarashi. “Ninja Cursors: Using Multiple Cursors to Assist Target Acquisition on Large Screens”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’08. Florence, Italy: Association for Computing Machinery, 2008, pp. 949–958. ISBN: 9781605580111. DOI: `10.1145/1357054.1357201` (cit. on p. 42).

[78] Werner A. König, Jens Gerken, Stefan Dierdorf, Harald Reiterer. “Adaptive Pointing – Design and Evaluation of a Precision Enhancing Technique for Absolute Pointing Devices”. In: *Human-Computer Interaction – INTERACT 2009*. Springer Berlin Heidelberg, 2009, pp. 658–671. DOI: `10.1007/978-3-642-03655-2_73` (cit. on pp. 17, 27, 30, 31, 34, 36, 38, 40, 44, 45).

[79] Regis Kopper, Felipe Bacim, Doug A. Bowman. “Rapid and accurate 3D selection by progressive refinement”. In: *2011 IEEE Symposium on 3D User Interfaces*. 3DUI ’11. IEEE, 2011, pp. 67–74. DOI: `10.1109/3DUI.2011.5759219` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 41, 44, 45).

[80] Regis Kopper, Doug A. Bowman, Mara G. Silva, Ryan P. McMahan. “A human motor behavior model for distal pointing tasks”. In: *International Journal of Human-Computer Studies* 68.10 (Oct. 2010), pp. 603–615. DOI: `10.1016/j.ijhcs.2010.05.001` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[81] Ernst Kruijff. “3D user interfaces: theory and practice”. In: *IEEE transactions on visualization and computer graphics* 18.4 (2012), pp. 565–572 (cit. on p. 16).

[82] Andre Kunert, Alexander Kulik, Anke Huckauf, Bernd Froehlich. “A Comparison of Tracking- and Controller-Based Input for Complex Bimanual Interaction in Virtual Environments”. In: *Eurographics Symposium on Virtual Environments*. Ed. by Bernd Froehlich, Roland Blach, Robert van Liere. The Eurographics Association, 2007. ISBN: 978-3-905674-02-6. DOI: `10.2312/EGVE/IPT_EGVE2007/043-052` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[83] Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, Mark Billinghurst. “Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–14. ISBN: 9781450356206. DOI: `10.1145/3173574.3173655` (cit. on pp. 27, 30–32, 34, 36, 38, 40, 44, 45).

[84] Jihyun Lee, Byungmoon Kim, Bongwon Suh, Eunyee Koh. “Exploring the Front Touch Interface for Virtual Reality Headsets”. In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’16. San Jose, California, USA: Association for Computing Machinery, 2016, pp. 2585–2591. ISBN: 9781450340823. DOI: `10.1145/2851581.2892344` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[85] Joong-Jae Lee, Jung-Min Park. “3D Mirrored Object Selection for Occluded Objects in Virtual Environments”. In: *IEEE Access* 8 (2020), pp. 200259–200274. ISSN: 2169-3536. DOI: `10.1109/ACCESS.2020.3035376` (cit. on pp. 13, 27, 30, 31, 34, 36, 38, 40–42, 44, 45).

[86] SangYoon Lee, Jinseok Seo, Gerard Jounghyun Kim, Chan-Mo Park. “Evaluation of pointing techniques for ray casting selection in virtual environments”. In: *Third International Conference on Virtual Reality and Its Application in Industry*. Ed. by Zhigeng Pan, Jiaoying Shi. Vol. 4756. International Society for Optics and Photonics. SPIE, 2003, pp. 38–44. DOI: `10.1117/12.497665` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[87] Cheng Lei, Kejimu Sunzi, Fengling Dai, Xiaoqin Liu, Yanfen Wang, Baolu Zhang, Lin He, Mei Ju. “Effects of virtual reality rehabilitation training on gait and balance in patients with Parkinson’s disease: A systematic review”. In: *PLoS One* 14.11 (2019). DOI: `10.1371/journal.pone.0224819` (cit. on p. 18).

[88] Jialei Li, Isaac Cho, Zachary Wartell. “Evaluation of Cursor Offset on 3D Selection in VR”. In: *Proceedings of the Symposium on Spatial User Interaction*. SUI ’18. Berlin, Germany: Association for Computing Machinery, 2018, pp. 120–129. ISBN: 9781450357081. DOI: `10.1145/3267782.3267797` (cit. on pp. 25, 27, 30, 31, 34–38, 42–45).

[89] Nianlong Li, Teng Han, Feng Tian, Jin Huang, Minghui Sun, Pourang Irani, Jason Alexander. “Get a Grip: Evaluating Grip Gestures for VR Input Using a Lightweight Pen”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI ’20. Honolulu, HI, USA: Association for Computing Machinery, 2020, pp. 1–13. ISBN: 9781450367080. DOI: `10.1145/3313831.3376698` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[90] Nianlong Li, Zhengquan Zhang, Can Liu, Zengyao Yang, Yinan Fu, Feng Tian, Teng Han, Mingming Fan. “VMirror: Enhancing the Interaction with Occluded or Distant Objects in VR with Virtual Mirrors”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2021. ISBN: 9781450380966. DOI: `10.1145/3411764.3445537` (cit. on pp. 27, 30, 31, 34, 36, 38, 40–42, 44, 45).

[91] Jiandong Liang, Mark Green. “JDCAD: A highly interactive 3D modeling system”. In: *Computers & Graphics* 18.4 (July 1994), pp. 499–506. DOI: `10.1016/0097-8493(94)90062-0` (cit. on pp. 15, 16, 27, 30, 31, 34, 36, 38, 44, 45).

[92] Zhuorui Liang, Xiangmin Xu, Chunmei Qing. “Natural Interaction in Freehand Pointing: Designing the mapping of interaction space.” In: *IEEE Consumer Electronics Magazine* 6.2 (2017), pp. 89–93. DOI: `10.1109/MCE.2016.2640740` (cit. on pp. 35, 37, 42).

[93] Chiuhsiang Joe Lin, Sui-Hua Ho, Yan-Jyun Chen. “An investigation of pointing postures in a 3D stereoscopic environment”. In: *Applied Ergonomics* 48 (May 2015), pp. 154–163. DOI: `10.1016/j.apergo.2014.12.001` (cit. on pp. 27, 30, 31, 34, 36–38, 44, 45).

[94] Nicholas G. Lipari, Christoph W. Borst. “Handymenu: Integrating menu selection into a multifunction smartphone-based VR controller”. In: *2015 IEEE Symposium on 3D User Interfaces*. 3DUI ’15. IEEE, Mar. 2015, pp. 129–132. DOI: `10.1109/3DUI.2015.7131737` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[95] Yingzhen Liu, Gang Wan. “Techniques for Selecting and Manipulating Object in Virtual Environment Based on 3-DOF Trackers and Data Glove”. In: *16th International Conference on Artificial Reality and Telexistence–Workshops*. ICAT ’06. IEEE, Nov. 2006, pp. 662–665. DOI: `10.1109/ICAT.2006.123` (cit. on pp. 35, 42, 44).

[96] Julian Looser, Mark Billinghurst, Raphael Grasset, Andy Cockburn. “An Evaluation of Virtual Lenses for Object Selection in Augmented Reality”. In: *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia*. GRAPHITE ’07. Perth, Australia: Association for Computing Machinery, 2007, pp. 203–210. ISBN: 9781595939128. DOI: `10.1145/1321261.1321297` (cit. on pp. 27, 30, 31, 34, 36–38, 44, 45).

[97] Xiaolong Lou, Andol X. Li, Ren Peng, Preben Hansen. “Optimising Free Hand Selection in Large Displays by Adapting to User’s Physical Movements”. In: *Proceedings of the 2016 Symposium on Spatial User Interaction*. SUI ’16. Tokyo, Japan: Association for Computing Machinery, 2016, pp. 23–31. ISBN: 9781450340687. DOI: `10.1145/2983310.2985754` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[98] Xueshi Lu, Difeng Yu, Hai-Ning Liang, Wenge Xu, Yuzheng Chen, Xiang Li, Khalad Hasan. “Exploration of Hands-free Text Entry Techniques For Virtual Reality”. In: *2020 IEEE International Symposium on Mixed and Augmented Reality*. ISMAR ’20. IEEE, Nov. 2020, pp. 344–349. DOI: `10.1109/ISMAR50242.2020.00061` (cit. on pp. 27, 30, 31, 34, 36, 38, 43–45).

[99] Yiqin Lu, Chun Yu, Yuanchun Shi. “Investigating Bubble Mechanism for Ray-Casting to Improve 3D Target Acquisition in Virtual Reality”. In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces*. VR ’20. IEEE, Mar. 2020, pp. 35–43. DOI: `10.1109/VR46266.2020.00021` (cit. on pp. 26, 27, 30, 31, 34, 36, 38, 40, 41, 44, 45).

[100] Siqi Luo, Robert J. Teather. “Camera-Based Selection with Cardboard HMDs”. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces*. VR ’19. IEEE, Mar. 2019, pp. 1066–1067. DOI: `10.1109/VR.2019.8797772` (cit. on pp. 27, 30–32, 34, 36–38, 43–45).

[101] Ville Mäkelä, Tomi Heimonen, Markku Turunen. “Magnetic Cursor: Improving Target Selection in Freehand Pointing Interfaces”. In: *Proceedings of The International Symposium on Pervasive Displays*. PerDis ’14. Copenhagen, Denmark: Association for Computing Machinery, 2014, pp. 112–117. ISBN: 9781450329521. DOI: `10.1145/2611009.2611025` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[102] Fabrice Matulic, Daniel Vogel. “Multiray: Multi-Finger Raycasting for Large Displays”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–13. ISBN: 9781450356206. DOI: `10.1145/3173574.3173819` (cit. on pp. 27, 30, 31, 34, 36–38, 44, 45).

[103] Sven Mayer, Jens Reinhardt, Robin Schweigert, Brighten Jelke, Valentin Schwind, Katrin Wolf, Niels Henze. “Improving Humans’ Ability to Interpret Deictic Gestures in Virtual Reality”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–14. ISBN: 9781450367080 (cit. on p. 18).

[104] Sven Mayer, Valentin Schwind, Robin Schweigert, Niels Henze. “The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–13. ISBN: 9781450356206. DOI: `10.1145/3173574.3174227` (cit. on p. 34).

[105] Sven Mayer, Katrin Wolf, Stefan Schneegass, Niels Henze. “Modeling Distant Pointing for Compensating Systematic Displacements”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI ’15. New York, NY, USA: Association for Computing Machinery, 2015, pp. 4165–4168. ISBN: 9781450331456 (cit. on p. 34).

[106] Daniel Medeiros, Lucas Teixeira, Felipe Carvalho, Ismael Santos, Alberto Raposo. “A Tablet-Based 3D Interaction Tool for Virtual Engineering Environments”. In: *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*. VRCAI ’13. Hong Kong, Hong Kong: Association for Computing Machinery, 2013, pp. 211–218. ISBN: 9781450325905. DOI: `10.1145/2534329.2534349` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[107] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, J. Jorge. “A Survey on 3D Virtual Object Manipulation: From the Desktop to Immersive Virtual Environments”. In: *Computer Graphics Forum* 38.1 (2019), pp. 21–45. DOI: `10.1111/cgf.13390` (cit. on p. 13).

[108] Daniel Mendes, Daniel Medeiros, Mauricio Sousa, Eduardo Cordeiro, Alfredo Ferreira, Joaquim A. Jorge. “Design and evaluation of a novel out-of-reach selection technique for VR using iterative refinement”. In: *Computers & Graphics* 67 (Oct. 2017), pp. 95–102. DOI: `10.1016/j.cag.2017.06.003` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 41, 43–45).

[109] Paul Milgram, Fumio Kishino. “A taxonomy of mixed reality visual displays”. In: *IEICE TRANSACTIONS on Information and Systems* 77.12 (1994), pp. 1321–1329 (cit. on p. 18).

[110] Patrick Millais, Simon L. Jones, Ryan Kelly. “Exploring Data in Virtual Reality: Comparisons with 2D Data Visualizations”. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA ’18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–6. ISBN: 9781450356213. DOI: `10.1145/3170427.3188537` (cit. on p. 18).

[111] Alexander Miloff, Philip Lindner, Peter Dafgård, Stefan Deak, Maria Garke, William Hamilton, Julia Heinsoo, Glenn Kristoffersson, Jonas Rafi, Kerstin Sindemark, Jessica Sjölund, Maria Zenger, Lena Reuterskiöld, Gerhard Andersson, Per Carlbring. “Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: A randomized non-inferiority trial”. In: *Behaviour Research and Therapy* 118 (2019), pp. 130–140. ISSN: 0005-7967. DOI: `https://doi.org/10.1016/j.brat.2019.04.004` (cit. on p. 18).

[112] M. Mine. “Virtual Environment Interaction Techniques”. In: 1995 (cit. on p. 16).

[113] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman. “Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement”. In: *BMJ* 339 (2009). DOI: `10.1136/bmj.b2535` (cit. on pp. 14, 21, 51, 53).

[114] Roberto A. Montano-Murillo, Cuong Nguyen, Rubaiat Habib Kazi, Sriram Subramanian, Stephen DiVerdi, Diego Martinez-Plasencia. “Slicing-Volume: Hybrid 3D/2D Multi-target Selection Technique for Dense Virtual Environments”. In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces*. VR ’20. IEEE, Mar. 2020, pp. 53–62. DOI: `10.1109/VR46266.2020.00023` (cit. on pp. 27, 30, 31, 34, 36, 38, 40–42, 44–46).

[115] Pedro Monteiro, Guilherme Gonçalves, Hugo Coelho, Miguel Melo, Maximino Bessa. “Hands-free interaction in immersive virtual reality: A systematic review”. In: *IEEE Transactions on Visualization and Computer Graphics*. TVGC ’21 27.5 (2021), pp. 2702–2713. DOI: `10.1109/TVCG.2021.3067687` (cit. on p. 21).

[116] Annette Mossel, Benjamin Venditti, Hannes Kaufmann. "DrillSample: Precise Selection in Dense Handheld Augmented Reality Environments". In: *Proceedings of the Virtual Reality International Conference: Laval Virtual*. VRIC '13. Laval, France: Association for Computing Machinery, 2013. ISBN: 9781450318754. DOI: `10.1145/2466816.2466827` (cit. on pp. 27, 30, 31, 34, 36, 38, 40–42, 44, 45).

[117] Joschka Mütterlein. "The three pillars of virtual reality? Investigating the roles of immersion, presence, and interactivity". In: *Proceedings of the 51st Hawaii international conference on system sciences*. 2018 (cit. on p. 51).

[118] Anh Nguyen, Amy Banic. "3DTouch: A wearable 3D input device for 3D applications". In: *2015 IEEE Virtual Reality*. VR '15. IEEE, Mar. 2015, pp. 55–61. DOI: `10.1109/VR.2015.7223324` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[119] Tomi Nukarinen, Jari Kangas, Jussi Rantala, Olli Koskinen, Roope Raisamo. "Evaluating Ray Casting and Two Gaze-Based Pointing Techniques for Object Selection in Virtual Reality". In: *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. VRST '18. Tokyo, Japan: Association for Computing Machinery, 2018. ISBN: 9781450360869. DOI: `10.1145/3281505.3283382` (cit. on pp. 13, 27, 30–32, 34, 36, 38, 44, 45).

[120] Dan R. Olsen, Travis Nielsen. "Laser Pointer Interaction". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '01. Seattle, Washington, USA: Association for Computing Machinery, 2001, pp. 17–22. ISBN: 1581133278. DOI: `10.1145/365024.365030` (cit. on p. 34).

[121] A. Olwal, H. Benko, S. Feiner. "SenseShapes: using statistical geometry for object selection in a multimodal augmented reality". In: *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings*. ISMAR '03. IEEE, Oct. 2003, pp. 300–301. DOI: `10.1109/ISMAR.2003.1240730` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[122] A. Olwal, S. Feiner. "The Flexible Pointer: An Interaction Technique for Selection in Augmented and Virtual Reality". In: 2003 (cit. on pp. 27, 30–32, 34, 36, 38, 40–42, 44, 45).

[123] Kasım Özacar, Juan David Hincapié-Ramos, Kazuki Takashima, Yoshifumi Kitamura. "3D Selection Techniques for Mobile Augmented Reality Head-Mounted Displays". In: *Interacting with Computers* 29.4 (July 2017), pp. 579–591. ISSN: 1873-7951. DOI: `10.1093/iwc/iww035` (cit. on pp. 27, 30–32, 34–38, 40, 44, 45).

[124] Hyung Min Park, Seok Han Lee, Jong Soo Choi. "Wearable augmented reality system using gaze interaction". In: *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. ISMAR '08. USA: IEEE Computer Society, Sept. 2008, pp. 175–176. ISBN: 9781424428403. DOI: `10.1109/ISMAR.2008.4637353` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[125] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, Hans Gellersen. "Gaze + Pinch Interaction in Virtual Reality". In: *Proceedings of the 5th Symposium on Spatial User Interaction*. SUI '17. Brighton, United Kingdom: Association for Computing Machinery, 2017, pp. 99–108. ISBN: 9781450354868. DOI: `10.1145/3131277.3132180` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 43–45).

[126] Duc-Minh Pham, Wolfgang Stuerzlinger. "Is the Pen Mightier than the Controller? A Comparison of Input Devices for Selection in Virtual and Augmented Reality". In: *25th ACM Symposium on Virtual Reality Software and Technology*. VRST '19. Parramatta, NSW, Australia: Association for Computing Machinery, 2019. ISBN: 9781450370011. DOI: `10.1145/3359996.3364264` (cit. on pp. 27, 29–31, 34–36, 38, 44, 45).

[127] Jeffrey S. Pierce, Andrew S. Forsberg, Matthew J. Conway, Seung Hong, Robert C. Zeleznik, Mark R. Mine. "Image Plane Interaction Techniques in 3D Immersive Environments". In: *Proceedings of the 1997 Symposium on Interactive 3D Graphics*. I3D '97. Providence, Rhode Island, USA: Association for Computing Machinery, 1997, 39–ff. ISBN: 0897918843. DOI: `10.1145/253284.253303` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[128] Krzysztof Pietroszek, Anastasia Kuzminykh, James R. Wallace, Edward Lank. "Smartcasting: A Discount 3D Interaction Technique for Public Displays". In: *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*. OzCHI '14. Sydney, New South Wales, Australia: Association for Computing Machinery, 2014, pp. 119–128. ISBN: 9781450306539. DOI: `10.1145/2686612.2686629` (cit. on pp. 30, 42, 44).

[129] Krzysztof Pietroszek, Liudmila Tahai, James R. Wallace, Edward Lank. "Watchcasting: Freehand 3D interaction with off-the-shelf smartwatch". In: *2017 IEEE Symposium on 3D User Interfaces*. 3DUI '17. IEEE, Mar. 2017, pp. 172–175. DOI: `10.1109/3DUI.2017.7893335` (cit. on pp. 27, 30, 31, 33, 34, 36, 38, 41, 44, 45).

[130] Krzysztof Pietroszek, James R. Wallace, Edward Lank. "Tiltcasting: 3D Interaction on Large Displays Using a Mobile Device". In: *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. UIST '15. Charlotte, NC, USA: Association for Computing Machinery, 2015, pp. 57–62. ISBN: 9781450337793. DOI: `10.1145/2807442.2807471` (cit. on pp. 27, 30, 31, 34, 36, 38, 41, 42, 44, 45).

[131] Barry A. Po, Brian D. Fisher, Kellogg S. Booth. "Comparing Cursor Orientations for Mouse, Pointer, and Pen Interaction". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '05. Portland, Oregon, USA: Association for Computing Machinery, 2005, pp. 291–300. ISBN: 1581139985. DOI: `10.1145/1054972.1055013` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[132] I. Poupyrev, T. Ichikawa, S. Weghorst, M. Billinghurst. "Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques". In: *Computer Graphics Forum* 17.3 (1998), pp. 41–52. DOI: `https://doi.org/10.1111/1467-8659.00252` (cit. on p. 16).

[133] Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, Tadao Ichikawa. "The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR". In: *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*. UIST '96. Seattle, Washington, USA: Association for Computing Machinery, 1996, pp. 79–80. ISBN: 0897917987. DOI: `10.1145/237091.237102` (cit. on pp. 27, 30, 31, 34–36, 38, 39, 41, 42, 44, 45).

[134] Ivan Poupyrev, Tadao Ichikawa. "Manipulating objects in virtual worlds: Categorization and empirical evaluation of interaction techniques". In: *Journal of Visual Languages & Computing* 10.1 (1999), pp. 19–35 (cit. on p. 16).

[135] Yuan Yuan Qian, Robert J. Teather. “The Eyes Don’t Have It: An Empirical Comparison of Head-Based and Eye-Based Selection in Virtual Reality”. In: *Proceedings of the 5th Symposium on Spatial User Interaction*. SUI ’17. Brighton, United Kingdom: Association for Computing Machinery, 2017, pp. 91–98. ISBN: 9781450354868. DOI: 10.1145/3131277.3132182 (cit. on pp. 25, 27, 30–32, 34, 36, 38, 44, 45).

[136] Vijay Rajanna, John Paulin Hansen. “Gaze Typing in Virtual Reality: Impact of Keyboard Design, Selection Method, and Motion”. In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. ETRA ’18. Warsaw, Poland: Association for Computing Machinery, 2018. ISBN: 9781450357067. DOI: 10.1145/3204493.3204541 (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[137] Adrian Ramcharitar, Robert J. Teather. “EZCursorVR: 2D Selection with Virtual Reality Head-Mounted Displays”. In: *Proceedings of the 44th Graphics Interface Conference*. GI ’18. Toronto, Canada: Canadian Human-Computer Communications Society, 2018, pp. 123–130. ISBN: 9780994786838. DOI: 10.20380/GI2018.17 (cit. on pp. 27, 30–32, 34–38, 44, 45).

[138] Umar Rashid, Aaron Quigley, Jarmo Kauko. “Selecting Targets on Large Display with Mobile Pointer and Touchscreen”. In: *ACM International Conference on Interactive Tabletops and Surfaces*. ITS ’10. Saarbrücken, Germany: Association for Computing Machinery, 2010, pp. 285–286. ISBN: 9781450303996. DOI: 10.1145/1936652.1936718 (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[139] Michael Rohs, Antti Oulasvirta. “Target Acquisition with Camera Phones When Used as Magic Lenses”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’08. Florence, Italy: Association for Computing Machinery, 2008, pp. 1409–1418. ISBN: 9781605580111. DOI: 10.1145/1357054.1357275 (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[140] Jannick P. Rolland, Richard L. Holloway, Henry Fuchs. “Comparison of optical and video see-through, head-mounted displays”. In: *Telemanipulator and Telepresence Technologies*. Ed. by Hari Das. Vol. 2351. International Society for Optics and Photonics. SPIE, 1995, pp. 293–307 (cit. on p. 18).

[141] Andrés Santos-Torres, Telmo Zarraonandia, Paloma Díaz, Ignacio Aedo. “Exploring Interaction Mechanisms for Map Interfaces in Virtual Reality Environments”. In: *Proceedings of the XIX International Conference on Human Computer Interaction*. Interacción 2018. Palma, Spain: Association for Computing Machinery, 2018. ISBN: 9781450364911. DOI: 10.1145/3233824.3233828 (cit. on pp. 27, 30–32, 34, 36, 38, 44, 45).

[142] Jonas Schjerlund, Kasper Hornbæk, Joanna Bergström. “Ninja Hands: Using Many Hands to Improve Target Selection in VR”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2021. ISBN: 9781450380966. DOI: 10.1145/3411764.3445759 (cit. on pp. 27, 30, 31, 34, 36, 38–42, 44, 45).

[143] G. Schmidt, Y. Baillot, D.G. Brown, E.B. Tomlin, J.E. Swan. “Toward Disambiguating Multiple Selections for Frustum-Based Pointing”. In: *3D User Interfaces*. 3DUI ’06. IEEE, Mar. 2006, pp. 87–94. DOI: 10.1109/VR.2006.133 (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[144] Robin Schweigert, Valentin Schwind, Sven Mayer. “EyePointing: A Gaze-Based Selection Technique”. In: *Proceedings of Mensch Und Computer 2019*. MuC’19. Hamburg, Germany: Association for Computing Machinery, 2019, pp. 719–723. ISBN: 9781450371988. DOI: `10.1145/3340764.3344897` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[145] Valentin Schwind, Sven Mayer, Alexandre Comeau-Vermeersch, Robin Schweigert, Niels Henze. “Up to the Finger Tip: The Effect of Avatars on Mid-Air Pointing Accuracy in Virtual Reality”. In: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY ’18. Melbourne, VIC, Australia: Association for Computing Machinery, 2018, pp. 477–488. ISBN: 9781450356244. DOI: `10.1145/3242671.3242675` (cit. on p. 35).

[146] Jakub Segen, Senthil Kumar. “Gesture VR: Vision-Based 3D Hand Interace for Spatial Interaction”. In: *Proceedings of the Sixth ACM International Conference on Multimedia*. MULTIMEDIA ’98. Bristol, United Kingdom: Association for Computing Machinery, 1998, pp. 455–464. ISBN: 0201309904. DOI: `10.1145/290747.290822` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[147] Zouhair Serrar, Nabil Elmarzouqi, Zahi Jarir, Jean-Christophe Lapayre. “Evaluation of Disambiguation Mechanisms of Object-Based Selection in Virtual Environment: Which Performances and Features to Support "Pick Out"?” In: *Proceedings of the XV International Conference on Human Computer Interaction*. Interacción ’14. Puerto de la Cruz, Tenerife, Spain: Association for Computing Machinery, 2014. ISBN: 9781450328807. DOI: `10.1145/2662253.2662282` (cit. on p. 39).

[148] Shaishav Siddhpuria, Sylvain Malacria, Mathieu Nancel, Edward Lank. “Pointing at a Distance with Everyday Smart Devices”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–11. ISBN: 9781450356206. DOI: `10.1145/3173574.3173747` (cit. on pp. 27, 30, 31, 33–38, 40, 44, 45).

[149] Ludwig Sidenmark, Christopher Clarke, Xuesong Zhang, Jenny Phu, Hans Gellersen. “Outline Pursuits: Gaze-Assisted Selection of Occluded Objects in Virtual Reality”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI ’20. Honolulu, HI, USA: Association for Computing Machinery, 2020, pp. 1–13. ISBN: 9781450367080. DOI: `10.1145/3313831.3376438` (cit. on pp. 27, 30–32, 34, 36, 38, 40, 41, 43–45).

[150] Ludwig Sidenmark, Hans Gellersen. “Eye & Head: Synergetic Eye and Head Movement for Gaze Pointing and Selection”. In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. UIST ’19. New Orleans, LA, USA: Association for Computing Machinery, 2019, pp. 1161–1174. ISBN: 9781450368162. DOI: `10.1145/3332165.3347921` (cit. on pp. 13, 27, 30–32, 34, 36, 38, 43–45).

[151] Ludwig Sidenmark, Dominic Potts, Bill Bapisch, Hans Gellersen. “Radi-Eye: Hands-Free Radial Interfaces for 3D Interaction Using Gaze-Activated Head-Crossing”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2021. ISBN: 9781450380966. DOI: `10.1145/3411764.3445697` (cit. on pp. 13, 27, 30, 31, 34, 36, 38, 44, 45).

[152] Jaisie Sin, Cosmin Munteanu. “Let’s Go There: Combining Voice and Pointing in VR”. In: *Proceedings of the 2nd Conference on Conversational User Interfaces*. CUI ’20. Bilbao, Spain: Association for Computing Machinery, 2020. ISBN: 9781450375443. DOI: 10.1145/3405755.3406161 (cit. on pp. 13, 27, 30, 31, 34, 36, 38, 40, 42, 44, 45).

[153] Mel Slater. “Immersion and the illusion of presence in virtual reality”. In: *British Journal of Psychology* 109.3 (2018), pp. 431–433. DOI: 10.1111/bjop.12305 (cit. on p. 13).

[154] Mel Slater, Vasilis Linakis, Martin Usoh, Rob Kooper. “Immersion, Presence and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess”. In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. VRST ’96. Hong Kong: Association for Computing Machinery, 1996, pp. 163–172. ISBN: 0897918258. DOI: 10.1145/3304181.3304216 (cit. on p. 51).

[155] Leonardo P. Soares, Soraia R. Musse, Mááarcio S. Pinho, Jean B. Boussu. “Evaluation of Selection Techniques on a Mobile Augmented Reality Game”. In: *2018 17th Brazilian Symposium on Computer Games and Digital Entertainment*. SBGames ’18. IEEE, Oct. 2018, pp. 127–12709. DOI: 10.1109/SBGAMES.2018.00024 (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[156] Chang Geun Song, No Jun Kwak, Dong Hyun Jeong. “Developing an Efficient Technique of Selection and Manipulation in Immersive V.E.” In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. VRST ’00. Seoul, Korea: Association for Computing Machinery, 2000, pp. 142–146. ISBN: 1581133162. DOI: 10.1145/502390.502417 (cit. on pp. 27, 30, 31, 34–36, 38, 41, 42, 44, 45).

[157] Marco Speicher, Anna Maria Feit, Pascal Ziegler, Antonio Krüger. “Selection-Based Text Entry in Virtual Reality”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–13. ISBN: 9781450356206. DOI: 10.1145/3173574.3174221 (cit. on pp. 26, 27, 30–32, 34–36, 38, 44, 45).

[158] A. Steed, C. Parker. “3D Selection Strategies for Head Tracked and Non-Head Tracked Operation of Spatially Immersive Displays”. In: 2004 (cit. on pp. 27, 30, 31, 34, 36, 38, 40–42, 44, 45).

[159] Frank Steinicke, Timo Ropinski, Klaus Hinrichs. “Object Selection in Virtual Environments Using an Improved Virtual Pointer Metaphor”. In: *Computational Imaging and Vision*. Kluwer Academic Publishers, pp. 320–326. DOI: 10.1007/1-4020-4179-9_46 (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[160] Rasmus Stenholt. “Efficient Selection of Multiple Objects on a Large Scale”. In: *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology*. VRST ’12. Toronto, Ontario, Canada: Association for Computing Machinery, 2012, pp. 105–112. ISBN: 9781450314695. DOI: 10.1145/2407336.2407357 (cit. on pp. 27, 30–32, 34, 36, 38, 40, 41, 44, 45).

[161] Richard Stoakley, Matthew J. Conway, Randy Pausch. “Virtual Reality on a WIM: Interactive Worlds in Miniature”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’95. Denver, Colorado, USA: ACM Press/Addison-Wesley Publishing Co., 1995, pp. 265–272. ISBN: 0201847051. DOI: 10.1145/223904.223938 (cit. on pp. 27, 30, 31, 34, 36, 38, 40–42, 44, 45).

[162] Minghui Sun, Mingming Cao, Limin Wang, Qian Qian. “PhoneCursor: Improving 3D Selection Performance With Mobile Device in AR”. In: *IEEE Access* 8 (2020), pp. 70616–70626. ISSN: 2169-3536. DOI: `10.1109/ACCESS.2020.2986037` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 41, 44, 45).

[163] Ivan Sutherland. “The ultimate display”. In: *Proceedings of the IFIPS Congress*. CUMINCAD, 1965 (cit. on p. 15).

[164] John Sweetser, Anders Grunnet-Jepsen, Gopal Panchanathan. “Absolute Pointing and Tracking Based Remote Control for Interactive User Experience”. In: UXTV ’08. Silicon Valley, California, USA: Association for Computing Machinery, 2008, pp. 155–164. ISBN: 9781605581002. DOI: `10.1145/1453805.1453837` (cit. on p. 34).

[165] Slawomir Konrad Tadeja, Timoleon Kipouros, Per Ola Kristensson. “Exploring Parallel Coordinates Plots in Virtual Reality”. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI EA ’19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–6. ISBN: 9781450359719. DOI: `10.1145/3290607.3313068` (cit. on p. 18).

[166] Tomomi Takashina, Mitsuru Ito, Hitoshi Nagaura, Eisuke Wakabayashi. “Evaluation of Curved Raycasting-based Interactive Surfaces in Virtual Environments”. In: *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. VRW ’21. IEEE, Mar. 2021, pp. 534–535. DOI: `10.1109/VRW52623.2021.00149` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[167] Vildan Tanriverdi, Robert J. K. Jacob. “Interacting with Eye Movements in Virtual Environments”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’00. The Hague, The Netherlands: Association for Computing Machinery, 2000, pp. 265–272. ISBN: 1581132166. DOI: `10.1145/332040.332443` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[168] Da Tao, Xiaofeng Diao, Tieyan Wang, Jingya Guo, Xingda Qu. “Freehand interaction with large displays: Effects of body posture, interaction distance and target size on task performance, perceived usability and workload”. In: *Applied Ergonomics* 93 (May 2021), p. 103370. DOI: `10.1016/j.apergo.2021.103370` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[169] Robert J. Teather, Wolfgang Stuerzlinger, Andriy Pavlovych. “Fishtank Fitts: A Desktop VR Testbed for Evaluating 3D Pointing Techniques”. In: *CHI ’14 Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’14. Toronto, Ontario, Canada: Association for Computing Machinery, 2014, pp. 519–522. ISBN: 9781450324748. DOI: `10.1145/2559206.2574810` (cit. on pp. 25, 27, 30, 31, 34, 36, 38, 44, 45).

[170] Bruce H. Thomas. “Evaluation of three input techniques for selection and annotation of physical objects through an augmented reality view”. In: *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*. ISMAR ’06. USA: IEEE Computer Society, Oct. 2006, pp. 33–36. ISBN: 1424406501. DOI: `10.1109/ISMAR.2006.297791` (cit. on pp. 27, 30–32, 34, 36, 38, 44, 45).

[171] Eleftherios Triantafyllidis, Zhibin Li. “The Challenges in Modeling Human Performance in 3D Space with Fitts’ Law”. In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2021. ISBN: 9781450380959 (cit. on p. 17).

[172] Huawei Tu, Susu Huang, Jiabin Yuan, Xiangshi Ren, Feng Tian. "Crossing-Based Selection with Virtual Reality Head-Mounted Displays". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland UK: Association for Computing Machinery, 2019, pp. 1–14. ISBN: 9781450359702. DOI: `10.1145/3290605.3300848` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[173] Stephen Uzor, Per Ola Kristensson. "An Exploration of Freehand Crossing Selection in Head-Mounted Augmented Reality". In: *ACM Trans. Comput.-Hum. Interact.* 28.5 (Aug. 2021). ISSN: 1073-0516. DOI: `10.1145/3462546` (cit. on p. 44).

[174] Lode Vanacken, Tovi Grossman, Karin Coninx. "Exploring the Effects of Environment Density and Target Visibility on Object Selection in 3D Virtual Environments". In: *2007 IEEE Symposium on 3D User Interfaces*. 3DUI '07. IEEE, Mar. 2007. DOI: `10.1109/3DUI.2007.340783` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 41, 44, 45).

[175] Lode Vanacken, Tovi Grossman, Karin Coninx. "Multimodal selection techniques for dense and occluded 3D virtual environments". In: *International Journal of Human-Computer Studies* 67.3 (Mar. 2009), pp. 237–255. DOI: `10.1016/j.ijhcs.2008.09.001` (cit. on pp. 27, 30, 31, 34, 36, 38, 40–42, 44–46).

[176] Rishi Vanukuru, Amarnath Murugan, Jayesh Pillai. "Dual Phone AR: Using a Second Phone as a Controller for Mobile Augmented Reality". In: *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology*. UIST '20 Adjunct. Virtual Event, USA: Association for Computing Machinery, 2020, pp. 117–119. ISBN: 9781450375153. DOI: `10.1145/3379350.3416139` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[177] Thomas Vincent, Laurence Nigay, Takeshi Kurata. "Handheld Augmented Reality: Effect of Registration Jitter on Cursor-Based Pointing Techniques". In: *Proceedings of the 25th Conference on l'Interaction Homme-Machine*. IHM '13. Talence, France: Association for Computing Machinery, 2013, pp. 1–6. ISBN: 9781450324076. DOI: `10.1145/2534903.2534905` (cit. on pp. 27, 30, 31, 34–36, 38, 44, 45).

[178] Daniel Vogel, Ravin Balakrishnan. "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*. UIST '05. Seattle, WA, USA: Association for Computing Machinery, 2005, pp. 33–42. ISBN: 1595932712. DOI: `10.1145/1095034.1095041` (cit. on pp. 21, 27, 30, 31, 34–38, 44, 45).

[179] Philipp Wacker, Oliver Nowak, Simon Voelker, Jan Borchers. "ARPen: Mid-Air Object Manipulation Techniques for a Bimanual AR System with Pen & Smartphone". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–12. ISBN: 9781450359702. DOI: `10.1145/3290605.3300849` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[180] Lili Wang, Jianjun Chen, Qixiang Ma, Voicu Popescu. "Disocclusion Headlight for Selection Assistance in VR". In: *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. VR '21. IEEE, Mar. 2021, pp. 216–225. DOI: `10.1109/VR50410.2021.00043` (cit. on pp. 27, 30, 31, 34, 36, 38, 40–42, 44, 45).

[181] Zhimin Wang, Huangyue Yu, Haofei Wang, Zongji Wang, Feng Lu. “Comparing Single-modal and Multimodal Interaction in an Augmented Reality System”. In: *2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct*. ISMAR Adjunct ’20. IEEE, Nov. 2020, pp. 165–166. DOI: `10.1109/ISMAR-Adjunct51615.2020.00052` (cit. on pp. 13, 27, 30, 31, 34, 36–38, 44, 45).

[182] Colin Ware, Kathy Lowther. “Selection Using a One-Eyed Cursor in a Fish Tank VR Environment”. In: *ACM Trans. Comput.-Hum. Interact.* 4.4 (Dec. 1997), pp. 309–322. ISSN: 1073-0516. DOI: `10.1145/267135.267136` (cit. on pp. 27, 30, 31, 34, 36, 38, 41, 44, 45).

[183] Matthias Weise, Raphael Zender, Ulrike Lucke. “A Comprehensive Classification of 3D Selection and Manipulation Techniques”. In: *Proceedings of Mensch Und Computer 2019*. MuC’19. Hamburg, Germany: Association for Computing Machinery, 2019, pp. 321–332. ISBN: 9781450371988. DOI: `10.1145/3340764.3340777` (cit. on pp. 13, 16, 21, 34, 37).

[184] Matt Whitlock, Ethan Harnner, Jed R. Brubaker, Shaun Kane, Danielle Albers Szafir. “Interacting with Distant Objects in Augmented Reality”. In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces*. VR ’18. IEEE, Mar. 2018, pp. 41–48. DOI: `10.1109/VR.2018.8446381` (cit. on pp. 27, 30–32, 34, 36, 38, 44, 45).

[185] Brett Wilkinson, Paul Calder. “Investigating Touch Interactions for an Augmented World”. In: *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat*. OZCHI ’08. Cairns, Australia: Association for Computing Machinery, 2008, pp. 25–32. ISBN: 0980306345. DOI: `10.1145/1517744.1517752` (cit. on pp. 27, 30, 31, 34–38, 44, 45).

[186] Jonathan Wonner, Jerome Grosjean, Antonio Capobianco, Dominique Bechmann. “Bubble Bee, an Alternative to Arrow for Pointing out Directions”. In: *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*. VRST ’13. Singapore: Association for Computing Machinery, 2013, pp. 97–100. ISBN: 9781450323796. DOI: `10.1145/2503713.2503753` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[187] Jonathan Wonner, Jerome Grosjean, Antonio Capobianco, Dominique Bechmann. “Starfish: A Selection Technique for Dense Virtual Environments”. In: *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology*. VRST ’12. Toronto, Ontario, Canada: Association for Computing Machinery, 2012, pp. 101–104. ISBN: 9781450314695. DOI: `10.1145/2407336.2407356` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 41, 44, 45).

[188] H.P. Wyss, R. Blach, M. Bues. “iSith - Intersection-based Spatial Interaction for Two Hands”. In: *3D User Interfaces (3DUI’06)*. 3DUI ’06. IEEE, Mar. 2006, pp. 59–61. DOI: `10.1109/VR.2006.93` (cit. on pp. 27, 30–32, 34, 36, 38, 41, 42, 44, 45).

[189] Chongbin Xu, Mingquan Zhou, Dongdong Zhang, Wuyang Shui, Zhongke Wu. “Guidance Rays: 3D Object Selection Based on Multi-Ray in Dense Scenario”. In: *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*. VRCAI ’13. Hong Kong, Hong Kong: Association for Computing Machinery, 2013, pp. 91–100. ISBN: 9781450325905. DOI: `10.1145/2534329.2534347` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 44, 45).

[190] Wenge Xu, Hai-Ning Liang, Anqi He, Zifan Wang. “Pointing and Selection Methods for Text Entry in Augmented Reality Head Mounted Displays”. In: *2019 IEEE International Symposium on Mixed and Augmented Reality*. ISMAR ’19. IEEE, Oct. 2019, pp. 279–288. DOI: `10.1109/ISMAR.2019.00026` (cit. on pp. 27, 30–32, 34, 36, 38, 44, 45).

[191] Yukang Yan, Yingtian Shi, Chun Yu, Yuanchun Shi. “HeadCross: Exploring Head-Based Crossing Selection on Head-Mounted Displays”. In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4.1 (Mar. 2020). DOI: `10.1145/3380983` (cit. on pp. 27, 30, 31, 34, 36, 38, 43–45).

[192] Yukang Yan, Chun Yu, Xin Yi, Yuanchun Shi. “HeadGesture: Hands-Free Input Approach Leveraging Head Movements for HMD Devices”. In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2.4 (Dec. 2018). DOI: `10.1145/3287076` (cit. on pp. 27, 30, 31, 33, 34, 36, 38, 44, 45).

[193] Jibin Yin, Chengyao Fu, Xiangliang Zhang, Tao Liu. “Precise Target Selection Techniques in Handheld Augmented Reality Interfaces”. In: *IEEE Access* 7 (2019), pp. 17663–17674. ISSN: 2169-3536. DOI: `10.1109/ACCESS.2019.2895219` (cit. on pp. 27, 30, 31, 34–36, 38, 40–42, 44, 45).

[194] Difeng Yu, Hai-Ning Liang, Xueshi Lu, Kaixuan Fan, Barrett Ens. “Modeling Endpoint Distribution of Pointing Selection Tasks in Virtual Reality Environments”. In: *ACM Trans. Graph.* 38.6 (Nov. 2019). ISSN: 0730-0301. DOI: `10.1145/3355089.3356544` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[195] Difeng Yu, Qiushi Zhou, Joshua Newn, Tilman Dingler, Eduardo Velloso, Jorge Goncalves. “Fully-Occluded Target Selection in Virtual Reality”. In: *IEEE Transactions on Visualization and Computer Graphics*. TVCG ’20 26.12 (Dec. 2020), pp. 3402–3413. ISSN: 1941-0506. DOI: `10.1109/TVCG.2020.3023606` (cit. on pp. 26, 27, 30, 31, 34, 36, 38, 40–42, 44, 45).

[196] Ja Eun Yu, Gerard J. Kim. “Blurry (Sticky) Finger: Proprioceptive Pointing and Selection of Distant Objects for Optical See-Through Based Augmented Reality”. In: *2016 IEEE International Symposium on Mixed and Augmented Reality*. ISMAR Adjunct ’16. IEEE, Sept. 2016, pp. 336–337. DOI: `10.1109/ISMAR-Adjunct.2016.0109` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[197] Anjie Zhu, Shiwei Cheng, Jing Fan. “Eye Tracking and Gesture Based Interaction for Target Selection on Large Displays”. In: *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. UbiComp ’18. Singapore: Association for Computing Machinery, 2018, pp. 319–322. ISBN: 9781450359665. DOI: `10.1145/3267305.3267607` (cit. on pp. 27, 30, 31, 34, 36, 38, 40, 41, 44, 45).

[198] Daniel Zielasko, Neha Neha, Benjamin Weyers, Torsten W. Kuhlen. “A reliable non-verbal vocal input metaphor for clicking”. In: *2017 IEEE Symposium on 3D User Interfaces*. 3DUI ’17. IEEE, Mar. 2017, pp. 40–49. DOI: `10.1109/3DUI.2017.7893316` (cit. on pp. 27, 30, 31, 34, 36, 38, 44, 45).

[199] David Zielinski, Brendan Macdonald, Regis Kopper. “Comparative study of input devices for a VR mine simulation”. In: *2014 IEEE Virtual Reality*. VR ’14. IEEE, Mar. 2014, pp. 125–126. DOI: `10.1109/VR.2014.6802083` (cit. on pp. 27, 30, 31, 34–38, 44, 45).

[200] Thomas G. Zimmerman, Jaron Lanier, Chuck Blanchard, Steve Bryson, Young Harvill. “A Hand Gesture Interface Device”. In: *SIGCHI Bull.* 18.4 (May 1986), pp. 189–192. ISSN: 0736-6906. DOI: `10.1145/1165387.275628` (cit. on p. 15).

All links were last followed on September 3, 2021.

Declaration

I hereby declare that the work presented in this thesis is entirely my own and that I did not use any other sources and references than the listed ones. I have marked all direct or indirect statements from other sources contained therein as quotations. Neither this work nor significant parts of it were part of another examination procedure. I have not published this work in whole or in part before. The electronic copy is consistent with all submitted copies.

place, date, signature